# **PCT**

# WORLD INTELLECTUAL PROPERTY ORGANIZATION International Bureau



# INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

	HED U	INDER THE PATENT COOPERATION TREATY (PCT)
(51) International Patent Classification 7:	, ,	(11) International Publication Number: WO 00/39303
C12N 15/49, C07K 14/16	A2	(43) International Publication Date: 6 July 2000 (06.07.00
(21) International Application Number: PCT/US (22) International Filing Date: 30 December 1999 ( (30) Priority Data: 60/114,495 31 December 1998 (31.12.9 60/156,670 29 September 1999 (29.09.9)	30.12.9 8) U	BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GE
(71) Applicant: CHIRON CORPORATION [US/US]; 456 Street, Emeryville, CA 94608 (US).	60 Hort	
(72) Inventors: BARNETT, Susan; Chiron Corporation, 4 ton Street - R440, Emeryville, CA 94608 (US). H Karin; Chiron Corporation, 4560 Horton Street Emeryville, CA 94608 (US). MARTIN, Eric; Ch poration, 4560 Horton Street - R440, Emeryville, (US).	IARTO R44 iron Co	<ul> <li>Published</li> <li>Without international search report and to be republished</li> <li>upon receipt of that report.</li> </ul>
(74) Agents: DOLLARD, Anne, S.; Chiron Corporation, tual Property - R440, P.O. Box 8097, Emery 94662-8097 (US) et al.	, Intelle ville, C	c- A
)		
(54) Title: MODIFIED HIV ENV POLYPEPTIDES		•
(57) Abstract		
Polynucleotide encoding modified HIV Env polypept of the CD4 binding region. Methods of diagnosis, treatme	ides are nt and	disclosed. The Env polypeptides are modified so as to expose at least pa prevention using the polynucleotides and polypeptides are also provided

# FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
A.M	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
ΑT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
ΑZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GB	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav	TM	Turkmenistan
BF	Burkina Faso	GR	Greece		Republic of Macedonia	TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin .	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Кепуа	NL	Netherlands	ΥU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	zw	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's	NZ	New Zealand		
CM	Cameroon		Republic of Korea	PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SID	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

# MODIFIED HIV ENV POLYPEPTIDES

# Technical Field

5

10

15

20

25

30

The invention relates generally to modified HIV envelope (Env) polypeptides which are useful as immunizing agents or for generating an immune response in a subject, for example a cellular immune response or a protective immune response. More particularly, the invention relates Env polypeptides such as gp120, gp140 or gp160, wherein at least one of the native  $\beta$ -sheet configurations has been modified. The invention also pertains to methods of using these polypeptides to elicit an immune response against a broad range of HIV subtypes.

# Background of the Invention

The human immunodeficiency virus (HIV-1, also referred to as HTLV-III, LAV or HTLV-III/LAV) is the etiological agent of the acquired immune deficiency syndrome (AIDS) and related disorders. (see, e.g., Barre-Sinoussi, et al., (1983) Science 220:868-871; Gallo et al. (1984) Science 224:500-503; Levy et al., (1984) Science 225:840-842; Siegal et al., (1981) N. Engl. J. Med. 305:1439-1444). AIDS patients usually have a long asymptomatic period followed by the progressive degeneration of the immune system and the central nervous system. Replication of the virus is highly regulated, and both latent and lytic infection of the CD4 positive helper subset of T-lymphocytes occur in tissue culture (Zagury et al., (1986) Science 231:850-853). Molecular studies of HIV-1 show that it encodes a number of genes (Ratner et al., (1985) Nature 313:277-284; Sanchez-Pescador et al., (1985) Science 227:484-492), including three structural genes -- gag, pol and env -- that are common to all retroviruses. Nucleotide sequences from viral genomes of other retroviruses, particularly HIV-2 and simian immunodeficiency viruses, SIV (previously referred to as STLV-III), also contain these structural genes. (Guyader et al., (1987) Nature 326:662-669; Chakrabarti et al., (1987) Nature

The envelope protein of HIV-1, HIV-2 and SIV is a glycoprotein of about 160 kd (gp160). During virus infection of the host cell, gp160 is cleaved by host cell proteases to form gp120 and the integral membrane protein, gp41. The gp41 portion is anchored in the

5

10

15

20

25

30

membrane bilayer of virion, while the gp120 segment protrudes into the surrounding environment. gp120 and gp41 are more covalently associated and free gp120 can be released from the surface of virions and infected cells.

As depicted in Figure 1, crystallography studies of the gp120 core polypeptide indicate that this polypeptide is folded into two major domains having certain emanating structures. The inner domain (inner with respect to the N and C terminus) features a two-helix, two-stranded bundle with a small five-stranded  $\beta$ -sandwich at its termini-proximal end and a projection at the distal end from which the V1/V2 stem emanates. The outer domain is a staked double barrel that lies along side the inner domain so that the outer barrel and inner bundle axes are approximately parallel. Between the distal inner domain and the distal outer domain is a four-stranded bridging sheet which holds a peculiar minidomain in contact with, but distinct from, the inner, the outer domain, and the V1/V2 domain. The bridging sheet is composed of four  $\beta$ -strand structures ( $\beta$ -3,  $\beta$ -2,  $\beta$ -21,  $\beta$ -20, shown in Figure 1). The bridging region can be seen in Figure 1 packing primarily over the inner domain, although some surface residues of the outer domain, such as Phe 382, reach into the bridging sheet to form part of its hydrophobic core.

The basic unit of the  $\beta$ -sheet conformation of the bridging sheet region is the  $\beta$ -strand which exists as a less tightly coiled helix, with 2.0 residues per turn. The  $\beta$ -strand conformation is only stable when incorporated into a  $\beta$ -sheet, where hydrogen bonds with close to optimal geometry are formed between the peptide groups on adjacent  $\beta$ -strands; the dipole moments of the strands are also aligned favorably. Side chains from adjacent residues of the same strand protrude from opposite sides of the sheet and do not interact with each other, but have significant interactions with their backbone and with the side chains of neighboring strands. For a general description of  $\beta$ -sheets, see, e.g., T.E. Creighton, Proteins: Structures and Molecular Properties (W.H. Freeman and Company, 1993); and A.L. Lehninger, Biochemistry (Worth Publishers, Inc., 1975).

The gp120 polypeptide is instrumental in mediating entry into the host cell. Recent studies have indicated that binding of CD4 to gp120 induces a conformational change in Env that allows for binding to a co-receptor (e.g., a chemokine receptor) and subsequent entry of the virus into the cell. (Wyatt, R., et al. (1998) Nature 393:705-711; Kwong, P., et al. (1998) Nature 393:648-659). Referring again to Figure 1, CD4 is bound into a depression formed at the interface of the outer domain, the inner domain and the bridging sheet of gp120.

Immunogenicity of the gp120 polypeptide has also been studied. For example, individuals infected by HIV-1 usually develop antibodies that can neutralize the virus in *in vitro* assays, and this response is directed primarily against linear neutralizing determinants in the third variable loop of gp120 glycoprotein (Javaherian, K., et al. (1989) *Proc. Natl. Acad. Sci.* 86:6786-6772; Matsushita, M., et al. (1988) *J. Virol.* 62:2107-2144; Putney, S., et al. (1986) *Science* 234:1392-1395; Rushe, J. R., et al. (1988) *Proc. Nat. Acad. Sci. USA* 85: 3198-3202.). However, these antibodies generally exhibit the ability to neutralize only a limited number of HIV-1 strains (Matthews, T. (1986) *Proc. Natl. Acad. Sci. USA*. 83:9709-9713; Nara, P. L., et al. (1988) *J. Virol.* 62:2622-2628; Palker, T. J., et al. (1988) *Proc. Natl. Acad. Sci. USA*. 85:1932-1936). Later in the course of HIV infection in humans, antibodies capable of neutralizing a wider range of HIV-1 isolates appear (Barre-Sinoussi, F., et al. (1983) *Science* 220:868-871; Robert-Guroff, M., et al. (1985) *Nature* (London) 316:72-74; Weis, R., et al. (1985) *Nature* (London) 316:72-755).

10

15

20

25

30

Recent work done by Stamatatos et al (1998) AIDS Res Hum Retroviruses 14(13):1129-39, shows that a deletion of the variable region 2 from a HIV-1<sub>SF162</sub> virus, which utilizes the CCR-5 co-receptor for virus entry, rendered the virus highly susceptible to serum-mediated neutralization. This V2 deleted virus was also neutralized by sera obtained from patients infected not only with clade B HIV-1 isolates but also with clade A, C, D and F HIV-1 isolates. However, deletion of the variable region 1 had no effect. Deletion of the variable regions 1 and 2 from a LAI isolate HIV-I<sub>IIIB</sub> also increased the susceptibility to neutralization by monoclonal antibodies whose epitopes are located within the V3 loop, the CD4-binding site, and conserved gp120 regions (Wyatt, R., et al. (1995) J Virol. 69:5723-5733). Rabbit immunogenicity studies done with the HIV-1 virus with deletions in the V1/V2 and V3 region from the LAI strain, which uses the CXCR4 co-receptor for virus entry, showed no improvement in the ability of Env to raise neutralizing antibodies (Leu et al. (1998) AIDS Res. and Human Retroviruses. 14:151-155).

Further, a subset of the broadly reactive antibodies, found in most infected individuals, interferes with the binding of gp120 and CD4 (Kang, C.-Y., et al. (1991) *Proc. Natl. Acad. Sci. USA.* 88:6171-6175; McDougal, J. S., et al. (1986) *J. Immunol.* 137:2937-2944). Other antibodies are believed to bind to the chemokine receptor binding region after CD4 has bound to Env (Thali et al. (1993) *J. Virol.* 67:3978-3988). The fact that neutralizing

antibodies generated during the course of HIV infection do not provide permanent antiviral effect may in part be due to the generation of "neutralization escapes" virus mutants and to the general decline in the host immune system associated with pathogenesis. In contrast, the presence of pre-existing neutralizing antibodies upon initial HIV-1 exposure will likely have a protective effect.

5

10

15

20

25

30

It is widely thought that a successful vaccine should be able to induce a strong, broadly neutralizing antibody response against diverse HIV-1 strains (Montefiori and Evans (1999) AIDS Res. Hum. Ret. 15(8):689-698; Bolognesi, D.,P., et al. (1994) Ann. Int. Med. 8:603-611; Haynes, B., F., et al. (1996) Science; 271: 324-328.). Neutralizing antibodies, by attaching to the incoming virions, can reduce or even prevent their infectivity for target cells and prevent the cell-to-cell spread of virus in tissue culture (Hu et al. (1992) Science 255:456-459; Burton, D.,R. and Montefiori, D. (1997) AIDS 11(suppl. A): 587-598). However as described above, antibodies directed against gp120 do not generally exhibit broad antibody responses against different HIV strains.

Currently, the focus of vaccine development, from the perspective of humoral immunity, is on the neutralization of primary isolates that utilize the CCR5 chemokine coreceptor believed to be important in virus entry (Zhu, T., et al. (1993) *Science* 261:1179-1181; Fiore, J., et al. (1994) Virology; 204:297-303). These viruses are generally much more resistant to antibody neutralization than T-cell line adapted strains that use the CXCR4 coreceptor, although both can be neutralized *in vitro* by certain broadly and potent acting monoclonal antibodies, such as IgG1b12, 2G12 and 2F5 (Trkola, A., et al. (1995) *J. Virol*. 69:6609-6617; D'Sousa PM., et al (1997) *J. Infect. Dis.* 175:1062-1075). These monoclonal antibodies are directed to the CD4 binding site, a glycosylation site and to the gp41 fusion domain, respectively. The problem that remains, however, is that it is not known how to induce antibodies of the appropriate specificity by vaccination. Antibodies (Abs) elicited by gp120 glycoprotein from a given isolate are usually only able to neutralize closely related viruses generally from similar, usually from the same, HIV-1 subtype.

Despite the above approaches, there remains a need for Env antigens that can elicit an immunological response (e.g., neutralizing and/or protective antibodies) in a subject against multiple HIV strains and subtypes, for example when administered as a vaccine. The present invention solves these and other problems by providing modified Env polypeptides (e.g., gp120) to expose epitopes in or near the CD4 binding site.

## Summary of the Invention

5

10

15

20

25

30

In accordance with the present invention, modified HIV Env polypeptides are provided. In particular, deletions and/or mutations are made in one or more of the  $4-\beta$  antiparallel-bridging sheet in the HIV Env polypeptide. In this way, enough structure is left to allow correct folding of the polypeptide, for example of gp120, yet enough of the bridging sheet is removed to expose the CD4 groove, allowing an immune response to be generated against epitopes in or near the CD4 binding site of the Env polypeptide (e.g., gp120).

In one aspect, the invention includes a polynucleotide encoding a modified HIV Env polypeptide wherein the polypeptide has at least one modified (e.g., deleted or replaced) amino acid residue deleted in the region corresponding to residues 421 to 436 relative to HXB-2, for example the constructs depicted in Figures 6-29 (SEQ ID NOs:3 to 26). In certain embodiments, the polynucleotide also has the region corresponding to residues 124-198 of the polypeptide HXB-2 (e.g., V1/V2) deleted and at least one amino acid deleted or replaced in the regions corresponding to the residues 119 to 123 and 199 to 210, relative to HXB-2. In other embodiments, these polynucleotides encode Env polypeptides having at least one amino acid of the small loop of the bridging sheet (e.g., amino acid residues 427 to 429 relative to HXB-2) deleted or replaced. The amino acid sequences of the modified polypeptides encoded by the polynucleotides of the present invention can be based on any HIV variant, for example SF162.

In another aspect, the invention includes immunogenic modified HIV Env polypeptides having at least one modified (e.g., deleted or replaced) amino acid residue deleted in the region corresponding to residues 421 to 436 relative to HXB-2, for example a deletion or replacement of one amino acids in the small loop region (e.g., amino acid residues 427 to 429 relative to HXB-2). These polypeptides may have modifications (e.g., a deletion or a replacement) of at least one amino acid between about amino acid residue 420 and amino acid residue 436, relative to HXB-2 and, optionally, may have deletions or truncations of the V1 and/or V2 regions. The immunogenic, modified polypeptides of the present invention can be based on any HIV variant, for example SF162.

In another aspect, the invention includes a vaccine composition comprising any of the polynucleotides encoding modified Env polypeptides described above. Vaccine compositions comprising the modified Env polypeptides and, optionally, an adjuvant are also included in the invention.

In yet another aspect, the invention includes a method of inducing an immune response in subject comprising, administering one or more of the polynucleotides or constructs described above in an amount sufficient to induce an immune response in the subject. In certain embodiments, the method further comprises administering an adjuvant to the subject.

In another aspect, the invention includes a method of inducing an immune response in a subject comprising administering a composition comprising any of the modified Env polypeptides described above and an adjuvant. The composition is administered in an amount sufficient to induce an immune response in the subject.

In another aspect, the invention includes a method of inducing an immune response in a subject comprising

- (a) administering a first composition comprising any of the polynucleotides described above in a priming step and
- (b) administering a second composition comprising any of the modified Env polypeptides described above, as a booster, in an amount sufficient to induce an immune response in the subject. In certain embodiments, the first composition, the second composition or both the first and second compositions further comprise an adjuvant.

These and other embodiments of the subject invention will readily occur to those of skill in the art in light of the disclosure herein.

20

25

30

5

10

15

### Brief Description of the Drawings

Figure 1 is a schematic depiction of the tertiary structure of the HIV-1<sub>HXB-2</sub> Env gp120 polypeptide, as determined by crystallography studies.

Figures 2A-C depict alignment of the amino acid sequence of wild-type HIV-1<sub>HXB-2</sub> Env gp160 polypeptide (SEQ ID NO:1) with amino acid sequence of HIV variants SF162 (shown as "162") (SEQ ID NO:2), SF2, CM236 and US4. Arrows indicate the regions that are deleted or replaced in the modified polypeptides. Black dots indicate conserved cysteine residues. The star indicates the position of the last amino acid in gp120.

Figures 3A-J depict alignment of nucleotide sequences of polynucleotides encoding modified Env polypeptides having V1/V2 deletions. The unmodified amino acid residues encoded by these sequences correspond to wildtype SF162 residues but are numbered relative to HXB-2.

Figures 4A-M depict alignment of nucleotide sequences of polynucleotides encoding modified Env polypeptides having deletions or replacements in the small loop. The unmodified amino acid residues encoded by these sequences correspond to wildtype SF162 residues but are numbered relative to HXB-2.

Figures 5A-N depict alignment of nucleotide sequences of polynucleotides encoding modified Env polypeptides having both V1/V2 deletions and, in addition, deletions or replacements in the small loop. The unmodified amino acid residues encoded by these sequences correspond to wildtype SF162 residues but are numbered relative to HXB-2.

5

10

15

20

25

Figure 6 depicts the nucleotide sequence of the construct designated Val120-Ala204 (SEO ID NO:3).

Figure 7 depicts the nucleotide sequence of the construct designated Val120-Ile201 (SEQ ID NO:4).

Figure 8 depicts the nucleotide sequence of the construct designated Val120-Ile201B (SEQ ID NO:5).

Figure 9 depicts the nucleotide sequence of the construct designated Lys121-Val200 (SEQ ID NO:6).

Figure 10 depicts the nucleotide sequence of the construct designated Leu122-Ser199 (SEQ ID NO:7).

Figure 11 depicts the nucleotide sequence of the construct designated Val120-Thr202 (SEQ ID NO:8).

Figure 12 depicts the nucleotide sequence of the construct designated Trp427-Gly431 (SEO ID NO:9).

Figure 13 depicts the nucleotide sequence of the construct designated Arg426-Gly431 (SEQ ID NO:10).

Figure 14 depicts the nucleotide sequence of the construct designated Arg426-Gly431B (SEQ ID NO:11).

Figure 15 depicts the nucleotide sequence of the construct designated Arg426-Lys432 (SEQ ID NO:12).

Figure 16 depicts the nucleotide sequence of the construct designated Asn425-Lys432 (SEQ ID NO:13).

Figure 17 depicts the nucleotide sequence of the construct designated Ile424-Ala433 (SEQ ID NO:14).

Figure 18 depicts the nucleotide sequence of the construct designated Ile423-Met434 (SEQ ID NO:15).

Figure 19 depicts the nucleotide sequence of the construct designated Gln422-Tyr435 (SEQ ID NO:16).

Figure 20 depicts the nucleotide sequence of the construct designated Gln422-Tyr435B (SEQ ID NO:17).

Figure 21 depicts the nucleotide sequence of the construct designated Leu122-Ser199;Arg426-Gly431 (SEQ ID NO:18).

Figure 22 depicts the nucleotide sequence of the construct designated Leu122-Ser199; Arg426-Lys432 (SEQ ID NO:19).

Figure 23 depicts the nucleotide sequence of the construct designated Leu122-Ser199; Trp427-Gly431 (SEQ ID NO:20).

Figure 24 depicts the nucleotide sequence of the construct designated Lys121-Val200; Asn425-Lys432 (SEQ ID NO:21).

Figure 25 depicts the nucleotide sequence of the construct designated Val120-Ile201; Ile424-Ala433 (SEQ ID NO:22).

Figure 26 depicts the nucleotide sequence of the construct designated Val120-Ile201B; Ile424-Ala433 (SEQ ID NO:23).

Figure 27 depicts the nucleotide sequence of the construct designated Val120-Thr202; Ile424-Ala433 (SEQ ID NO:24).

Figure 28 depicts the nucleotide sequence of the construct designated Val127-Asn195 (SEO ID NO:25).

Figure 29 depicts the nucleotide sequence of the construct designated Val127-Asn195; Arg426-Gly431 (SEQ ID NO:26).

Detailed Description of the Invention

5

10

20

25

30

The practice of the present invention will employ, unless otherwise indicated, conventional methods of protein chemistry, viral immunobiology, molecular biology and recombinant DNA techniques within the skill of the art. Such techniques are explained fully in the literature. See, e.g., T.E. Creighton, <u>Proteins: Structures and Molecular Properties</u> (W.H. Freeman and Company, 1993); Nelson L.M. and Jerome H.K. <u>HIV Protocols</u> in Methods in Molecular Medicine, vol. 17, 1999; Sambrook, et al., <u>Molecular Cloning: A</u>

<u>Laboratory Manual</u> (Cold Spring Harbor Laboratory, 1989); F.M. Ausubel et al. <u>Current Protocols in Molecular Biology</u>, Greene Publishing Associates & Wiley Interscience New York; and Lipkowitz and Boyd, <u>Reviews in Computational Chemistry</u>, volumes 1-present (Wiley-VCH, New York, New York, 1999).

It must be noted that, as used in this specification and the appended claims, the singular forms "a", "an" and "the" include plural referents unless the content clearly dictates otherwise. Thus, for example, reference to "a polypeptide" includes a mixture of two or more polypeptides, and the like.

### 10 Definitions

5

15

20

25

30

In describing the present invention, the following terms will be employed, and are intended to be defined as indicated below.

The terms "polypeptide," and "protein" are used interchangeably herein to denote any polymer of amino acid residues. The terms encompass peptides, oligopeptides, dimers, multimers, and the like. Such polypeptides can be derived from natural sources or can be synthesized or recombinantly produced. The terms also include postexpression modifications of the polypeptide, for example, glycosylation, acetylation, phosphorylation, etc.

A polypeptide as defined herein is generally made up of the 20 natural amino acids Ala (A), Arg (R), Asn (N), Asp (D), Cys (C), Gln (Q), Glu (E), Gly (G), His (H), Ile (I), Leu (L), Lys (K), Met (M), Phe (F), Pro (P), Ser (S), Thr (T), Trp (W), Tyr (Y) and Val (V) and may also include any of the several known amino acid analogs, both naturally occurring and synthesized analogs, such as but not limited to homoisoleucine, asaleucine, 2-(methylenecyclopropyl)glycine, S-methylcysteine, S-(prop-1-enyl)cysteine, homoserine, ornithine, norleucine, norvaline, homoarginine, 3-(3-carboxyphenyl)alanine, cyclohexylalanine, mimosine, pipecolic acid, 4-methylglutamic acid, canavanine, 2,3-diaminopropionic acid, and the like. Further examples of polypeptide agents which will find use in the present invention are set forth below.

By "geometry" or "tertiary structure" of a polypeptide or protein is meant the overall 3-D configuration of the protein. As described herein, the geometry can be determined, for example, by crystallography studies or by using various programs or algorithms which predict the geometry based on interactions between the amino acids making up the primary and secondary structures.

By "wild type" polypeptide, polypeptide agent or polypeptide drug, is meant a naturally occurring polypeptide sequence, and its corresponding secondary structure. An "isolated" or "purified" protein or polypeptide is a protein which is separate and discrete from a whole organism with which the protein is normally associated in nature. It is apparent that the term denotes proteins of various levels of purity. Typically, a composition containing a purified protein will be one in which at least about 35%, preferably at least about 40-50%, more preferably, at least about 75-85%, and most preferably at least about 90% or more, of the total protein in the composition will be the protein in question.

5

10

15

20

25

30

By "Env polypeptide" is meant a molecule derived from an envelope protein, preferably from HIV Env. The envelope protein of HIV-1 is a glycoprotein of about 160 kd (gp160). During virus infection of the host cell, gp160 is cleaved by host cell proteases to form gp120 and the integral membrane protein, gp41. The gp41 portion is anchored in (and spans) the membrane bilayer of virion, while the gp120 segment protrudes into the surrounding environment. As there is no covalent attachment between gp120 and gp41, free gp120 is released from the surface of virions and infected cells. Env polypeptides may also include gp140 polypeptides. Env polypeptides can exist as monomers, dimers or multimers.

By a "gp120 polypeptide" is meant a molecule derived from a gp120 region of the Env polypeptide. Preferably, the gp120 polypeptide is derived from HIV Env. The primary amino acid sequence of gp120 is approximately 511 amino acids, with a polypeptide core of about 60,000 daltons. The polypeptide is extensively modified by N-linked glycosylation to increase the apparent molecular weight of the molecule to 120,000 daltons. The amino acid sequence of gp120 contains five relatively conserved domains interspersed with five hypervariable domains. The positions of the 18 cysteine residues in the gp120 primary sequence of the HIV-1<sub>HXB-2</sub> (hereinafter "HXB-2") strain, and the positions of 13 of the approximately 24 N-linked glycosylation sites in the gp120 sequence are common to most, if not all, gp120 sequences. The hypervariable domains contain extensive amino acid substitutions, insertions and deletions. Despite this variation, most, if not all, gp120 sequences preserve the virus's ability to bind to the viral receptor CD4. A "gp120 polypeptide" includes both single subunits or multimers.

Env polypeptides (e.g., gp120, gp140 and gp160) include a "bridging sheet" comprised of 4 anti-parallel  $\beta$ -strands ( $\beta$ -2,  $\beta$ -3,  $\beta$ -20 and  $\beta$ -21) that form a  $\beta$ -sheet. Extruding from one pair of the  $\beta$ -strands ( $\beta$ -2 and  $\beta$ -3) are two loops, V1 and V2. The  $\beta$ -2

sheet occurs at approximately amino acid residue 119 (Cys) to amino acid residue 123 (Thr) while β-3 occurs at approximately amino acid residue 199 (Ser) to amino acid residue 201 (Ile), relative to HXB-2. The "V1/V2 region" occurs at approximately amino acid positions 126 (Cys) to residue 196 (Cys), relative to HXB-2. (see, *e.g.*, Wyatt et al. (1995) *J. Virol.* 69:5723-5733; Stamatatos et al. (1998) *J. Virol.* 72:7840-7845). Extruding from the second pair of β-strands (β-20 and β-21) is a "small-loop" structure, also referred to herein as "the bridging sheet small loop." In HXB-2, β-20 extends from about amino acid residue 422 (Gln) to amino acid residue 426 (Met) while β-21 extends from about amino acid residue 430 (Val) to amino acid residue 435 (Tyr). In variant SF162, the Met-426 is an Arg (R) residue. The "small loop" extends from about amino acid residue 427 (Trp) through 429 (Lys), relative to HXB-2. A representative diagram of gp120 showing the bridging sheet, the small loop, and V1/V2 is shown in Figure 1. In addition, alignment of the amino acid sequences of Env polypeptide gp160 of selected variants is shown, relative to HXB-2, in Figures 2A-C.

5

10

15

20

25

30

Furthermore, an "Env polypeptide" or "gp120 polypeptide" as defined herein is not limited to a polypeptide having the exact sequence described herein. Indeed, the HIV genome is in a state of constant flux and contains several variable domains which exhibit relatively high degrees of variability between isolates. It is readily apparent that the terms encompass Env (e.g., gp120) polypeptides from any of the identified HIV isolates, as well as newly identified isolates, and subtypes of these isolates. Descriptions of structural features are given herein with reference to HXB-2. One of ordinary skill in the art in view of the teachings of the present disclosure and the art can determine corresponding regions in other HIV variants (e.g., isolates HIV<sub>IIIb</sub>, HIV<sub>SF2</sub>, HIV-1<sub>SF162</sub>, HIV-1<sub>SF170</sub>, HIV<sub>LAV</sub>, HIV<sub>LAV</sub>, HIV<sub>MN</sub>, HIV-1<sub>CM235</sub>, HIV-1<sub>US4</sub>, other HIV-1 strains from diverse subtypes(e.g., subtypes, A through G, and O), HIV-2 strains and diverse subtypes (e.g., HIV-2<sub>UC1</sub> and HIV-2<sub>UC2</sub>), and simian immunodeficiency virus (SIV). (See, e.g., Virology, 3rd Edition (W.K. Joklik ed. 1988); Fundamental Virology, 2nd Edition (B.N. Fields and D.M. Knipe, eds. 1991); Virology, 3rd Edition (Fields, BN, DM Knipe, PM Howley, Editors, 1996, Lippincott-Raven, Philadelphia, PA; for a description of these and other related viruses), using for example, sequence comparison programs (e.g., BLAST and others described herein) or identification and alignment of structural features (e.g., a program such as the "ALB" program described herein that can identify  $\beta$ -sheet regions). The actual amino acid sequences of the modified Env polypeptides can be based on any HIV variant.

Additionally, the term "Env polypeptide" (e.g., "gp120 polypeptide") encompasses proteins which include additional modifications to the native sequence, such as additional internal deletions, additions and substitutions. These modifications may be deliberate, as through site-directed mutagenesis, or may be accidental, such as through naturally occurring mutational events. Thus, for example, if the Env polypeptide is to be used in vaccine compositions, the modifications must be such that immunological activity (i.e., the ability to elicit an antibody response to the polypeptide) is not lost. Similarly, if the polypeptides are to be used for diagnostic purposes, such capability must be retained.

5

10

15

20

25

30

Thus, a "modified Env polypeptide" is an Env polypeptide (e.g., gp120 as defined above), which has been manipulated to delete or replace all or a part of the bridging sheet portion and, optionally, the variable regions V1 and V2. Generally, modified Env (e.g., gp120) polypeptides have enough of the bridging sheet removed to expose the CD4 binding site, but leave enough of the structure to allow correct folding (e.g., correct geometry). Thus, modifications to the  $\beta$ -20 and  $\beta$ -21 regions (between about amino acid residues 420 and 435 relative to HXB-2) are preferred. Additionally, modifications to the  $\beta$ -2 and  $\beta$ -3 regions (between about amino acid residues 119 (Cys) and 201 (Ile)) and modifications (e.g., truncations) to the V1 and V2 loop regions may also be made. Although not all possible  $\beta$ -sheet and V1/V2 modifications have been exemplified herein, it is to be understood that other disrupting modifications are also encompassed by the present invention.

Normally, such a modified polypeptide is capable of secretion into growth medium in which an organism expressing the protein is cultured. However, for purposes of the present invention, such polypeptides may also be recovered intracellularly. Secretion into growth media is readily determined using a number of detection techniques, including, e.g., polyacrylamide gel electrophoresis and the like, and immunological techniques such as Western blotting and immunoprecipitation assays as described in, e.g., International Publication No. WO 96/04301, published February 15, 1996.

A gp120 or other Env polypeptide is produced "intracellularly" when it is found within the cell, either associated with components of the cell, such as in association with the endoplasmic reticulum (ER) or the Golgi Apparatus, or when it is present in the soluble cellular fraction. The gp120 and other Env polypeptides of the present invention may also be secreted into growth medium so long as sufficient amounts of the polypeptides remain

present within the cell such that they can be purified from cell lysates using techniques described herein.

5

10

15

20

25

30

An "immunogenic" gp120 or other Env protein is a molecule that includes at least one epitope such that the molecule is capable of either eliciting an immunological reaction in an individual to which the protein is administered or, in the diagnostic context, is capable of reacting with antibodies directed against the HIV in question.

By "epitope" is meant a site on an antigen to which specific B cells and/or T cells respond, rendering the molecule including such an epitope capable of eliciting an immunological reaction or capable of reacting with HIV antibodies present in a biological sample. The term is also used interchangeably with "antigenic determinant" or "antigenic determinant site." An epitope can comprise 3 or more amino acids in a spatial conformation unique to the epitope. Generally, an epitope consists of at least 5 such amino acids and, more usually, consists of at least 8-10 such amino acids. Methods of determining spatial conformation of amino acids are known in the art and include, for example, x-ray crystallography and 2-dimensional nuclear magnetic resonance. Furthermore, the identification of epitopes in a given protein is readily accomplished using techniques well known in the art, such as by the use of hydrophobicity studies and by site-directed serology. See, also, Geysen et al., Proc. Natl. Acad. Sci. USA (1984) 81:3998-4002 (general method of rapidly synthesizing peptides to determine the location of immunogenic epitopes in a given antigen); U.S. Patent No. 4,708,871 (procedures for identifying and chemically synthesizing epitopes of antigens); and Geysen et al., Molecular Immunology (1986) 23:709-715 (technique for identifying peptides with high affinity for a given antibody). Antibodies that recognize the same epitope can be identified in a simple immunoassay showing the ability of one antibody to block the binding of another antibody to a target antigen.

An "immunological response" or "immune response" as used herein is the development in the subject of a humoral and/or a cellular immune response to the Env (e.g., gp120) polypeptide when the polypeptide is present in a vaccine composition. These antibodies may also neutralize infectivity, and/or mediate antibody-complement or antibody dependent cell cytotoxicity to provide protection to an immunized host. Immunological reactivity may be determined in standard immunoassays, such as a competition assays, well known in the art.

Techniques for determining amino acid sequence "similarity" are well known in the art. In general, "similarity" means the exact amino acid to amino acid comparison of two or more polypeptides at the appropriate place, where amino acids are identical or possess similar chemical and/or physical properties such as charge or hydrophobicity. A so-termed "percent similarity" then can be determined between the compared polypeptide sequences.

Techniques for determining nucleic acid and amino acid sequence identity also are well known in the art and include determining the nucleotide sequence of the mRNA for that gene (usually via a cDNA intermediate) and determining the amino acid sequence encoded thereby, and comparing this to a second amino acid sequence. In general, "identity" refers to an exact nucleotide to nucleotide or amino acid to amino acid correspondence of two polynucleotides or polypeptide sequences, respectively.

10

15

20

25

30

Two or more polynucleotide sequences can be compared by determining their "percent identity." Two or more amino acid sequences likewise can be compared by determining their "percent identity." The percent identity of two sequences, whether nucleic acid or peptide sequences, is generally described as the number of exact matches between two aligned sequences divided by the length of the shorter sequence and multiplied by 100. An approximate alignment for nucleic acid sequences is provided by the local homology algorithm of Smith and Waterman, Advances in Applied Mathematics 2:482-489 (1981). This algorithm can be extended to use with peptide sequences using the scoring matrix developed by Dayhoff, Atlas of Protein Sequences and Structure, M.O. Dayhoff ed., 5 suppl. 3:353-358, National Biomedical Research Foundation, Washington, D.C., USA, and normalized by Gribskov, Nucl. Acids Res. 14(6):6745-6763 (1986). An implementation of this algorithm for nucleic acid and peptide sequences is provided by the Genetics Computer Group (Madison, WI) in their BestFit utility application. The default parameters for this method are described in the Wisconsin Sequence Analysis Package Program Manual, Version 8 (1995) (available from Genetics Computer Group, Madison, WI). Other equally suitable programs for calculating the percent identity or similarity between sequences are generally known in the art.

For example, percent identity of a particular nucleotide sequence to a reference sequence can be determined using the homology algorithm of Smith and Waterman with a default scoring table and a gap penalty of six nucleotide positions. Another method of establishing percent identity in the context of the present invention is to use the MPSRCH

package of programs copyrighted by the University of Edinburgh, developed by John F. Collins and Shane S. Sturrok, and distributed by IntelliGenetics, Inc. (Mountain View, CA). From this suite of packages, the Smith-Waterman algorithm can be employed where default parameters are used for the scoring table (for example, gap open penalty of 12, gap extension penalty of one, and a gap of six). From the data generated, the "Match" value reflects "sequence identity." Other suitable programs for calculating the percent identity or similarity between sequences are generally known in the art, such as the alignment program BLAST, which can also be used with default parameters. For example, BLASTN and BLASTP can be used with the following default parameters: genetic code = standard; filter = none; strand = both; cutoff = 60; expect = 10; Matrix = BLOSUM62; Descriptions = 50 sequences; sort by = HIGH SCORE; Databases = non-redundant, GenBank + EMBL + DDBJ + PDB + GenBank CDS translations + Swiss protein + Spupdate + PIR. Details of these programs can be found at the following internet address: http://www.ncbi.nlm.gov/cgi-bin/BLAST.

One of skill in the art can readily determine the proper search parameters to use for a given sequence in the above programs. For example, the search parameters may vary based on the size of the sequence in question. Thus, for example, a representative embodiment of the present invention would include an isolated polynucleotide having X contiguous nucleotides, wherein (i) the X contiguous nucleotides have at least about 50% identity to Y contiguous nucleotides derived from any of the sequences described herein, (ii) X equals Y, and (iii) X is greater than or equal to 6 nucleotides and up to 5000 nucleotides, preferably greater than or equal to 8 nucleotides and up to 5000 nucleotides, more preferably 10-12 nucleotides and up to 5000 nucleotides, and even more preferably 15-20 nucleotides, up to the number of nucleotides present in the full-length sequences described herein (e.g., see the Sequence Listing and claims), including all integer values falling within the above-described ranges.

The synthetic expression cassettes (and purified polynucleotides) of the present invention include related polynucleotide sequences having about 80% to 100%, greater than 80-85%, preferably greater than 90-92%, more preferably greater than 95%, and most preferably greater than 98% sequence (including all integer values falling within these described ranges) identity to the synthetic expression cassette sequences disclosed herein (for example, to the claimed sequences or other sequences of the present invention) when the sequences of the present invention are used as the query sequence.

Computer programs are also available to determine the likelihood of certain polypeptides to form structures such as β-sheets. One such program, described herein, is the "ALB" program for protein and polypeptide secondary structure calculation and predication. In addition, secondary protein structure can be predicted from the primary amino acid sequence, for example using protein crystal structure and aligning the protein sequence related to the crystal structure (e.g., using Molecular Operating Environment (MOE) programs available from the Chemical Computing Group Inc., Montreal, P.Q., Canada). Other methods of predicting secondary structures are described, for example, in Garnier et al. (1996) Methods Enzymol. 266:540-553; Geourjon et al. (1995) Comput. Applic. Biosci. 11:681-684; Levin (1997) Protein Eng. 10:771-776; and Rost et al. (1993) J. Molec. Biol. 232:584-599.

Homology can also be determined by hybridization of polynucleotides under conditions which form stable duplexes between homologous regions, followed by digestion with single-stranded-specific nuclease(s), and size determination of the digested fragments. Two DNA, or two polypeptide sequences are "substantially homologous" to each other when the sequences exhibit at least about 80%-85%, preferably at least about 90%, and most preferably at least about 95%-98% sequence identity over a defined length of the molecules, as determined using the methods above. As used herein, substantially homologous also refers to sequences showing complete identity to the specified DNA or polypeptide sequence. DNA sequences that are substantially homologous can be identified in a Southern hybridization experiment under, for example, stringent conditions, as defined for that particular system. Defining appropriate hybridization conditions is within the skill of the art. See, e.g., Sambrook et al., supra; DNA Cloning, supra; Nucleic Acid Hybridization, supra.

A "coding sequence" or a sequence which "encodes" a selected protein, is a nucleic acid sequence which is transcribed (in the case of DNA) and translated (in the case of mRNA) into a polypeptide in vitro or in vivo when placed under the control of appropriate regulatory sequences. The boundaries of the coding sequence are determined by a start codon at the 5' (amino) terminus and a translation stop codon at the 3' (carboxy) terminus. A coding sequence can include, but is not limited to cDNA from viral nucleotide sequences as well as synthetic and semisynthetic DNA sequences and sequences including base analogs. A transcription termination sequence may be located 3' to the coding sequence.

"Control elements" refers collectively to promoter sequences, ribosome binding sites, polyadenylation signals, transcription termination sequences, upstream regulatory domains, enhancers, and the like, which collectively provide for the transcription and translation of a coding sequence in a host cell. Not all of these control elements need always be present so long as the desired gene is capable of being transcribed and translated.

5

10

15

20

25

30

A control element "directs the transcription" of a coding sequence in a cell when RNA polymerase will bind the promoter sequence and transcribe the coding sequence into mRNA, which is then translated into the polypeptide encoded by the coding sequence.

"Operably linked" refers to an arrangement of elements wherein the components so described are configured so as to perform their usual function. Thus, control elements operably linked to a coding sequence are capable of effecting the expression of the coding sequence when RNA polymerase is present. The control elements need not be contiguous with the coding sequence, so long as they function to direct the expression thereof. Thus, for example, intervening untranslated yet transcribed sequences can be present between, e.g., a promoter sequence and the coding sequence and the promoter sequence can still be considered "operably linked" to the coding sequence.

"Recombinant" as used herein to describe a nucleic acid molecule means a polynucleotide of genomic, cDNA, semisynthetic, or synthetic origin which, by virtue of its origin or manipulation: (1) is not associated with all or a portion of the polynucleotide with which it is associated in nature; and/or (2) is linked to a polynucleotide other than that to which it is linked in nature. The term "recombinant" as used with respect to a protein or polypeptide means a polypeptide produced by expression of a recombinant polynucleotide. "Recombinant host cells," "host cells," "cells," "cell lines," "cell cultures," and other such terms denoting procaryotic microorganisms or eucaryotic cell lines cultured as unicellular entities, are used interchangeably, and refer to cells which can be, or have been, used as recipients for recombinant vectors or other transfer DNA, and include the progeny of the original cell which has been transfected. It is understood that the progeny of a single parental cell may not necessarily be completely identical in morphology or in genomic or total DNA complement to the original parent, due to accidental or deliberate mutation. Progeny of the parental cell which are sufficiently similar to the parent to be characterized by the relevant property, such as the presence of a nucleotide sequence encoding a desired peptide, are included in the progeny intended by this definition, and are covered by the above terms.

By "vertebrate subject" is meant any member of the subphylum chordata, including, without limitation, humans and other primates, including non-human primates such as chimpanzees and other apes and monkey species; farm animals such as cattle, sheep, pigs, goats and horses; domestic mammals such as dogs and cats; laboratory animals including rodents such as mice, rats and guinea pigs; birds, including domestic, wild and game birds such as chickens, turkeys and other gallinaceous birds, ducks, geese, and the like. The term does not denote a particular age. Thus, both adult and newborn individuals are intended to be covered.

As used herein, a "biological sample" refers to a sample of tissue or fluid isolated from an individual, including but not limited to, for example, blood, plasma, serum, fecal matter, urine, bone marrow, bile, spinal fluid, lymph fluid, samples of the skin, external secretions of the skin, respiratory, intestinal, and genitourinary tracts, samples derived from the gastric epithelium and gastric mucosa, tears, saliva, milk, blood cells, organs, biopsies and also samples of *in vitro* cell culture constituents including but not limited to conditioned media resulting from the growth of cells and tissues in culture medium, e.g., recombinant cells, and cell components.

The terms "label" and "detectable label" refer to a molecule capable of detection, including, but not limited to, radioactive isotopes, fluorescers, chemiluminescers, enzymes, enzyme substrates, enzyme cofactors, enzyme inhibitors, chromophores, dyes, metal ions, metal sols, ligands (e.g., biotin or haptens) and the like. The term "fluorescer" refers to a substance or a portion thereof which is capable of exhibiting fluorescence in the detectable range. Particular examples of labels which may be used with the invention include, but are not limited to fluorescein, rhodamine, dansyl, umbelliferone, Texas red, luminol, acradimum esters, NADPH,  $\alpha$ - $\beta$ -galactosidase, horseradish peroxidase, glucose oxidase, alkaline phosphatase and urease.

#### Overview

5

10

15

20

25

30

The present invention concerns modified Env polypeptide molecules (e.g., glycoprotein ("gp") 120). Without being bound by a particular theory, it appears that it has been difficult to generate immunological responses against Env because the CD4 binding site is buried between the outer domain, the inner domain and the V1/V2 domains. Thus, although deletion of the V1/V2 domain may render the virus more susceptible to

neutralization by monoclonal antibody directed to the CD4 site, the bridging sheet covering most of the CD4 binding domain may prevent an antibody response. Thus, the present invention provides Env polypeptides that maintain their general overall structure yet expose the CD4 binding domain. This allows the generation of an immune response (e.g., an antibody response) to epitopes in or near the CD4 binding site.

Various forms of the different embodiments of the invention, described herein, may be combined.

## **β-Sheet Conformations**

5

10

15

20

25

30

In the present invention, location of the  $\beta$ -sheet structures were identified relative to 3-D (crystal) structure of an HXB-2 crystallized Env protein (see, Example 1A). Based on this structure, constructs encoding polypeptides having replacements and or excisions which maintain overall geometry while exposing the CD4 binding site were designed. In particular, the crystal structure of HXB-2 was downloaded from the Brookhaven Database. Using the default parameters of the Loop Search feature of the Biopolymer module of the Sybyl molecular modeling package, homology and fit of amino acids which could replace the native loops between  $\beta$ -strands yet maintain overall tertiary structure were determined. Constructs encoding the modified Env polypeptides were then designed (Example 1.B.).

Thus, the modified Env polypeptides typically have enough of the bridging sheet removed to expose the CD4 groove, but have enough of the structure to allow correct folding of the Env glycoprotein. Exemplary constructs are described below.

# **Polypeptide Production**

The polypeptides of the present invention can be produced in any number of ways which are well known in the art.

In one embodiment, the polypeptides are generated using recombinant techniques, well known in the art. In this regard, oligonucleotide probes can be devised based on the known sequences of the Env (e.g., gp120) polypeptide genome and used to probe genomic or cDNA libraries for Env genes. The gene can then be further isolated using standard techniques and, e.g., restriction enzymes employed to truncate the gene at desired portions of the full-length sequence. Similarly, the Env gene(s) can be isolated directly from cells and tissues containing the same, using known techniques, such as phenol extraction and the

sequence further manipulated to produce the desired truncations. See, e.g., Sambrook et al., supra, for a description of techniques used to obtain and isolate DNA.

The genes encoding the modified (e.g., truncated and/or substituted) polypeptides can be produced synthetically, based on the known sequences. The nucleotide sequence can be designed with the appropriate codons for the particular amino acid sequence desired. The complete sequence is generally assembled from overlapping oligonucleotides prepared by standard methods and assembled into a complete coding sequence. See, e.g., Edge (1981) Nature 292:756; Nambair et al. (1984) Science 223:1299; Jay et al. (1984) J. Biol. Chem. 259:6311; Stemmer et al. (1995) Gene 164:49-53.

5

10

15

20

25

30

Recombinant techniques are readily used to clone a gene encoding an Env polypeptide gene which can then be mutagenized *in vitro* by the replacement of the appropriate base pair(s) to result in the codon for the desired amino acid. Such a change can include as little as one base pair, effecting a change in a single amino acid, or can encompass several base pair changes. Alternatively, the mutations can be effected using a mismatched primer which hybridizes to the parent nucleotide sequence (generally cDNA corresponding to the RNA sequence), at a temperature below the melting temperature of the mismatched duplex. The primer can be made specific by keeping primer length and base composition within relatively narrow limits and by keeping the mutant base centrally located. See, *e.g.*, Innis et al, (1990) PCR Applications: Protocols for Functional Genomics; Zoller and Smith, *Methods Enzymol*. (1983) 100:468. Primer extension is effected using DNA polymerase, the product cloned and clones containing the mutated DNA, derived by segregation of the primer extended strand, selected. Selection can be accomplished using the mutant primer as a hybridization probe. The technique is also applicable for generating multiple point mutations. See, e.g., Dalbie-McFarland et al. *Proc. Natl. Acad. Sci USA* (1982) 79:6409.

Once coding sequences for the desired proteins have been isolated or synthesized, they can be cloned into any suitable vector or replicon for expression. As will be apparent from the teachings herein, a wide variety of vectors encoding modified polypeptides can be generated by creating expression constructs which operably link, in various combinations, polynucleotides encoding Env polypeptides having deletions or mutation therein. Thus, polynucleotides encoding a particular deleted V1/V2 region can be operably linked with polynucleotides encoding polypeptides having deletions or replacements in the small loop

region and the construct introduced into a host cell for polypeptide expression. Non-limiting examples of such combinations are discussed in the Examples.

Numerous cloning vectors are known to those of skill in the art, and the selection of an appropriate cloning vector is a matter of choice. Examples of recombinant DNA vectors for cloning and host cells which they can transform include the bacteriophage λ (E. coli), pBR322 (E. coli), pACYC177 (E. coli), pKT230 (gram-negative bacteria), pGV1106 (gram-negative bacteria), pLAFR1 (gram-negative bacteria), pME290 (non-E. coli gram-negative bacteria), pHV14 (E. coli and Bacillus subtilis), pBD9 (Bacillus), pIJ61 (Streptomyces), pUC6 (Streptomyces), YIp5 (Saccharomyces), YCp19 (Saccharomyces) and bovine papilloma virus (mammalian cells). See, generally, DNA Cloning: Vols. I & II, supra; Sambrook et al., supra; B. Perbal, supra.

5

10

15

20

25

30

Insect cell expression systems, such as baculovirus systems, can also be used and are known to those of skill in the art and described in, e.g., Summers and Smith, Texas Agricultural Experiment Station Bulletin No. 1555 (1987). Materials and methods for baculovirus/insect cell expression systems are commercially available in kit form from, inter alia, Invitrogen, San Diego CA ("MaxBac" kit).

Plant expression systems can also be used to produce the modified Env proteins. Generally, such systems use virus-based vectors to transfect plant cells with heterologous genes. For a description of such systems see, e.g., Porta et al., *Mol. Biotech.* (1996) 5:209-221; and Hackland et al., *Arch. Virol.* (1994) 139:1-22.

Viral systems, such as a vaccinia based infection/transfection system, as described in Tomei et al., *J. Virol.* (1993) 67:4017-4026 and Selby et al., *J. Gen. Virol.* (1993) 74:1103-1113, will also find use with the present invention. In this system, cells are first transfected *in vitro* with a vaccinia virus recombinant that encodes the bacteriophage T7 RNA polymerase. This polymerase displays exquisite specificity in that it only transcribes templates bearing T7 promoters. Following infection, cells are transfected with the DNA of interest, driven by a T7 promoter. The polymerase expressed in the cytoplasm from the vaccinia virus recombinant transcribes the transfected DNA into RNA which is then translated into protein by the host translational machinery. The method provides for high level, transient, cytoplasmic production of large quantities of RNA and its translation product(s).

The gene can be placed under the control of a promoter, ribosome binding site (for bacterial expression) and, optionally, an operator (collectively referred to herein as "control" elements), so that the DNA sequence encoding the desired Env polypeptide is transcribed into RNA in the host cell transformed by a vector containing this expression construction. The coding sequence may or may not contain a signal peptide or leader sequence. With the present invention, both the naturally occurring signal peptides or heterologous sequences can be used. Leader sequences can be removed by the host in post-translational processing. See, e.g., U.S. Patent Nos. 4,431,739; 4,425,437; 4,338,397. Such sequences include, but are not limited to, the TPA leader, as well as the honey bee mellitin signal sequence.

10

15

5

Other regulatory sequences may also be desirable which allow for regulation of expression of the protein sequences relative to the growth of the host cell. Such regulatory sequences are known to those of skill in the art, and examples include those which cause the expression of a gene to be turned on or off in response to a chemical or physical stimulus, including the presence of a regulatory compound. Other types of regulatory elements may also be present in the vector, for example, enhancer sequences.

The control sequences and other regulatory sequences may be ligated to the coding sequence prior to insertion into a vector. Alternatively, the coding sequence can be cloned directly into an expression vector which already contains the control sequences and an appropriate restriction site.

20

25

30

In some cases it may be necessary to modify the coding sequence so that it may be attached to the control sequences with the appropriate orientation; *i.e.*, to maintain the proper reading frame. Mutants or analogs may be prepared by the deletion of a portion of the sequence encoding the protein, by insertion of a sequence, and/or by substitution of one or more nucleotides within the sequence. Techniques for modifying nucleotide sequences, such as site-directed mutagenesis, are well known to those skilled in the art. See, e.g., Sambrook et al., supra; DNA Cloning, Vols. I and II, supra; Nucleic Acid Hybridization, supra.

The expression vector is then used to transform an appropriate host cell. A number of mammalian cell lines are known in the art and include immortalized cell lines available from the American Type Culture Collection (ATCC), such as, but not limited to, Chinese hamster ovary (CHO) cells, HeLa cells, baby hamster kidney (BHK) cells, monkey kidney cells (COS), human hepatocellular carcinoma cells (e.g., Hep G2), Vero293 cells, as well as others. Similarly, bacterial hosts such as *E. coli*, *Bacillus subtilis*, and *Streptococcus spp.*, will find

use with the present expression constructs. Yeast hosts useful in the present invention include inter alia, Saccharomyces cerevisiae, Candida albicans, Candida maltosa, Hansenula polymorpha, Kluyveromyces fragilis, Kluyveromyces lactis, Pichia guillerimondii, Pichia pastoris, Schizosaccharomyces pombe and Yarrowia lipolytica. Insect cells for use with baculovirus expression vectors include, inter alia, Aedes aegypti, Autographa californica, Bombyx mori, Drosophila melanogaster, Spodoptera frugiperda, and Trichoplusia ni.

5

10

15

20

25

30

Depending on the expression system and host selected, the proteins of the present invention are produced by growing host cells transformed by an expression vector described above under conditions whereby the protein of interest is expressed. The selection of the appropriate growth conditions is within the skill of the art.

In one embodiment, the transformed cells secrete the polypeptide product into the surrounding media. Certain regulatory sequences can be included in the vector to enhance secretion of the protein product, for example using a tissue plasminogen activator (TPA) leader sequence, a γ-interferon signal sequence or other signal peptide sequences from known secretory proteins. The secreted polypeptide product can then be isolated by various techniques described herein, for example, using standard purification techniques such as but not limited to, hydroxyapatite resins, column chromatography, ion-exchange chromatography, size-exclusion chromatography, electrophoresis, HPLC, immunoadsorbent techniques, affinity chromatography, immunoprecipitation, and the like..

Alternatively, the transformed cells are disrupted, using chemical, physical or mechanical means, which lyse the cells yet keep the Env polypeptides substantially intact. Intracellular proteins can also be obtained by removing components from the cell wall or membrane, e.g., by the use of detergents or organic solvents, such that leakage of the Env polypeptides occurs. Such methods are known to those of skill in the art and are described in, e.g., *Protein Purification Applications: A Practical Approach*, (E.L.V. Harris and S. Angal, Eds., 1990)

For example, methods of disrupting cells for use with the present invention include but are not limited to: sonication or ultrasonication; agitation; liquid or solid extrusion; heat treatment; freeze-thaw; desiccation; explosive decompression; osmotic shock; treatment with lytic enzymes including proteases such as trypsin, neuraminidase and lysozyme; alkali treatment; and the use of detergents and solvents such as bile salts, sodium dodecylsulphate,

Triton, NP40 and CHAPS. The particular technique used to disrupt the cells is largely a matter of choice and will depend on the cell type in which the polypeptide is expressed, culture conditions and any pre-treatment used.

5

10

15

20

25

30

Following disruption of the cells, cellular debris is removed, generally by centrifugation, and the intracellularly produced Env polypeptides are further purified, using standard purification techniques such as but not limited to, column chromatography, ion-exchange chromatography, size-exclusion chromatography, electrophoresis, HPLC, immunoadsorbent techniques, affinity chromatography, immunoprecipitation, and the like.

For example, one method for obtaining the intracellular Env polypeptides of the present invention involves affinity purification, such as by immunoaffinity chromatography using anti-Env specific antibodies, or by lectin affinity chromatography. Particularly preferred lectin resins are those that recognize mannose moieties such as but not limited to resins derived from Galanthus nivalis agglutinin (GNA), Lens culinaris agglutinin (LCA or lentil lectin), Pisum sativum agglutinin (PSA or pea lectin), Narcissus pseudonarcissus agglutinin (NPA) and Allium ursinum agglutinin (AUA). The choice of a suitable affinity resin is within the skill in the art. After affinity purification, the Env polypeptides can be further purified using conventional techniques well known in the art, such as by any of the techniques described above.

It may be desirable to produce Env (e.g., gp120) complexes, either with itself or other proteins. Such complexes are readily produced by e.g., co-transfecting host cells with constructs encoding for the Env (e.g., gp120) and/or other polypeptides of the desired complex. Co-transfection can be accomplished either in trans or cis, i.e., by using separate vectors or by using a single vector which bears both of the Env and other gene. If done using a single vector, both genes can be driven by a single set of control elements or, alternatively, the genes can be present on the vector in individual expression cassettes, driven by individual control elements. Following expression, the proteins will spontaneously associate.

Alternatively, the complexes can be formed by mixing the individual proteins together which have been produced separately, either in purified or semi-purified form, or even by mixing culture media in which host cells expressing the proteins, have been cultured. See, International Publication No. WO 96/04301, published February 15, 1996, for a description of such complexes.

5

10

15

20

25

30

Relatively small polypeptides, i.e., up to about 50 amino acids in length, can be conveniently synthesized chemically, for example by any of several techniques that are known to those skilled in the peptide art. In general, these methods employ the sequential addition of one or more amino acids to a growing peptide chain. Normally, either the amino or carboxyl group of the first amino acid is protected by a suitable protecting group. The protected or derivatized amino acid can then be either attached to an inert solid support or utilized in solution by adding the next amino acid in the sequence having the complementary (amino or carboxyl) group suitably protected, under conditions that allow for the formation of an amide linkage. The protecting group is then removed from the newly added amino acid residue and the next amino acid (suitably protected) is then added, and so forth. After the desired amino acids have been linked in the proper sequence, any remaining protecting groups (and any solid support, if solid phase synthesis techniques are used) are removed sequentially or concurrently, to render the final polypeptide. By simple modification of this general procedure, it is possible to add more than one amino acid at a time to a growing chain, for example, by coupling (under conditions which do not racemize chiral centers) a protected tripeptide with a properly protected dipeptide to form, after deprotection, a pentapeptide. See, e.g., J. M. Stewart and J. D. Young, Solid Phase Peptide Synthesis (Pierce Chemical Co., Rockford, IL 1984) and G. Barany and R. B. Merrifield, The Peptides: Analysis, Synthesis, Biology, editors E. Gross and J. Meienhofer, Vol. 2, (Academic Press, New York, 1980), pp. 3-254, for solid phase peptide synthesis techniques; and M. Bodansky, Principles of Peptide Synthesis, (Springer-Verlag, Berlin 1984) and E. Gross and J. Meienhofer, Eds., The Peptides: Analysis, Synthesis, Biology, Vol. 1, for classical solution synthesis.

Typical protecting groups include t-butyloxycarbonyl (Boc), 9-fluorenylmethoxycarbonyl (Fmoc) benzyloxycarbonyl (Cbz); p-toluenesulfonyl (Tx); 2,4-dinitrophenyl; benzyl (Bzl); biphenylisopropyloxycarboxy-carbonyl, t-amyloxycarbonyl, isobornyloxycarbonyl, o-bromobenzyloxycarbonyl, cyclohexyl, isopropyl, acetyl, o-nitrophenylsulfonyl and the like.

Typical solid supports are cross-linked polymeric supports. These can include divinylbenzene cross-linked-styrene-based polymers, for example, divinylbenzene-hydroxymethylstyrene copolymers, divinylbenzene-chloromethylstyrene copolymers and divinylbenzene-benzhydrylaminopolystyrene copolymers.

The polypeptide analogs of the present invention can also be chemically prepared by other methods such as by the method of simultaneous multiple peptide synthesis. See, e.g., Houghten *Proc. Natl. Acad. Sci. USA* (1985) 82:5131-5135; U.S. Patent No. 4,631,211.

# Diagnostic and Vaccine Applications

5

10

15

20

25

30

The intracellularly produced Env polypeptides of the present invention, complexes thereof, or the polynucleotides coding therefor, can be used for a number of diagnostic and therapeutic purposes. For example, the proteins and polynucleotides or antibodies generated against the same, can be used in a variety of assays, to determine the presence of reactive antibodies/and or Env proteins in a biological sample to aid in the diagnosis of HIV infection or disease status or as measure of response to immunization.

The presence of antibodies reactive with the Env (e.g., gp120) polypeptides and, conversely, antigens reactive with antibodies generated thereto, can be detected using standard electrophoretic and immunodiagnostic techniques, including immunoassays such as competition, direct reaction, or sandwich type assays. Such assays include, but are not limited to, western blots; agglutination tests; enzyme-labeled and mediated immunoassays, such as ELISAs; biotin/avidin type assays; radioimmunoassays; immunoelectrophoresis; immunoprecipitation, etc. The reactions generally include revealing labels such as fluorescent, chemiluminescent, radioactive, or enzymatic labels or dye molecules, or other methods for detecting the formation of a complex between the antigen and the antibody or antibodies reacted therewith.

Solid supports can be used in the assays such as nitrocellulose, in membrane or microtiter well form; polyvinylchloride, in sheets or microtiter wells; polystyrene latex, in beads or microtiter plates; polyvinylidine fluoride; diazotized paper; nylon membranes; activated beads, and the like.

Typically, the solid support is first reacted with the biological sample (or the gp120 proteins), washed and then the antibodies, (or a sample suspected of containing antibodies), applied. After washing to remove any non-bound ligand, a secondary binder moiety is added under suitable binding conditions, such that the secondary binder is capable of associating selectively with the bound ligand. The presence of the secondary binder can then be detected using techniques well known in the art. Typically, the secondary binder will comprise an antibody directed against the antibody ligands. A number of anti-human immunoglobulin

(Ig) molecules are known in the art (e.g., commercially available goat anti-human Ig or rabbit anti-human Ig). Ig molecules for use herein will preferably be of the IgG or IgA type, however, IgM may also be appropriate in some instances. The Ig molecules can be readily conjugated to a detectable enzyme label, such as horseradish peroxidase, glucose oxidase, Beta-galactosidase, alkaline phosphatase and urease, among others, using methods known to those of skill in the art. An appropriate enzyme substrate is then used to generate a detectable signal.

5

10

15

20

25

30

Alternatively, a "two antibody sandwich" assay can be used to detect the proteins of the present invention. In this technique, the solid support is reacted first with one or more of the antibodies directed against Env (e.g., gp120), washed and then exposed to the test sample. Antibodies are again added and the reaction visualized using either a direct color reaction or using a labeled second antibody, such as an anti-immunoglobulin labeled with horseradish peroxidase, alkaline phosphatase or urease.

Assays can also be conducted in solution, such that the viral proteins and antibodies thereto form complexes under precipitating conditions. The precipitated complexes can then be separated from the test sample, for example, by centrifugation. The reaction mixture can be analyzed to determine the presence or absence of antibody-antigen complexes using any of a number of standard methods, such as those immunodiagnostic methods described above.

The modified Env proteins, produced as described above, or antibodies to the proteins, can be provided in kits, with suitable instructions and other necessary reagents, in order to conduct immunoassays as described above. The kit can also contain, depending on the particular immunoassay used, suitable labels and other packaged reagents and materials (i.e. wash buffers and the like). Standard immunoassays, such as those described above, can be conducted using these kits.

The Env polypeptides and polynucleotides encoding the polypeptides can also be used in vaccine compositions, individually or in combination, in e.g., prophylactic (i.e., to prevent infection) or therapeutic (to treat HIV following infection) vaccines. The vaccines can comprise mixtures of one or more of the modified Env proteins (or nucleotide sequences encoding the proteins), such as Env (e.g., gp120) proteins derived from more than one viral isolate. The vaccine may also be administered in conjunction with other antigens and immunoregulatory agents, for example, immunoglobulins, cytokines, lymphokines, and chemokines, including but not limited to IL-2, modified IL-2 (cys125-ser125), GM-CSF, IL-

12, γ-interferon, IP-10, MIP1β and RANTES. The vaccines may be administered as polypeptides or, alternatively, as naked nucleic acid vaccines (e.g., DNA), using viral vectors (e.g., retroviral vectors, adenoviral vectors, adeno-associated viral vectors) or non-viral vectors (e.g., liposomes, particles coated with nucleic acid or protein). The vaccines may also comprise a mixture of protein and nucleic acid, which in turn may be delivered using the same or different vehicles. The vaccine may be given more than once (e.g., a "prime" administration followed by one or more "boosts") to achieve the desired effects. The same composition can be administered as the prime and as the one or more boosts. Alternatively, different compositions can be used for priming and boosting.

5

10

15

20

25

30

The vaccines will generally include one or more "pharmaceutically acceptable excipients or vehicles" such as water, saline, glycerol, ethanol, etc. Additionally, auxiliary substances, such as wetting or emulsifying agents, pH buffering substances, and the like, may be present in such vehicles.

A carrier is optionally present which is a molecule that does not itself induce the production of antibodies harmful to the individual receiving the composition. Suitable carriers are typically large, slowly metabolized macromolecules such as proteins, polysaccharides, polylactic acids, polyglycollic acids, polymeric amino acids, amino acid copolymers, lipid aggregates (such as oil droplets or liposomes), and inactive virus particles. Such carriers are well known to those of ordinary skill in the art. Furthermore, the Env polypeptide may be conjugated to a bacterial toxoid, such as toxoid from diphtheria, tetanus, cholera, etc.

Adjuvants may also be used to enhance the effectiveness of the vaccines. Such adjuvants include, but are not limited to: (1) aluminum salts (alum), such as aluminum hydroxide, aluminum phosphate, aluminum sulfate, etc.; (2) oil-in-water emulsion formulations (with or without other specific immunostimulating agents such as muramyl peptides (see below) or bacterial cell wall components), such as for example (a) MF59 (International Publication No. WO 90/14837), containing 5% Squalene, 0.5% Tween 80, and 0.5% Span 85 (optionally containing various amounts of MTP-PE (see below), although not required) formulated into submicron particles using a microfluidizer such as Model 110Y microfluidizer (Microfluidics, Newton, MA), (b) SAF, containing 10% Squalane, 0.4% Tween 80, 5% pluronic-blocked polymer L121, and thr-MDP (see below) either microfluidized into a submicron emulsion or vortexed to generate a larger particle size

5

10

15

20

25

30

emulsion, and (c) Ribi™ adjuvant system (RAS), (Ribi Immunochem, Hamilton, MT) containing 2% Squalene, 0.2% Tween 80, and one or more bacterial cell wall components from the group consisting of monophosphorylipid A (MPL), trehalose dimycolate (TDM), and cell wall skeleton (CWS), preferably MPL + CWS (Detox<sup>TM</sup>); (3) saponin adjuvants, such as Stimulon™ (Cambridge Bioscience, Worcester, MA) may be used or particle generated therefrom such as ISCOMs (immunostimulating complexes); (4) Complete Freunds Adjuvant (CFA) and Incomplete Freunds Adjuvant (IFA); (5) cytokines, such as interleukins (IL-1, IL-2, etc.), macrophage colony stimulating factor (M-CSF), tumor necrosis factor (TNF), etc.; (6) detoxified mutants of a bacterial ADP-ribosylating toxin such as a cholera toxin (CT), a pertussis toxin (PT), or an E. coli heat-labile toxin (LT), particularly LT-K63 (where lysine is substituted for the wild-type amino acid at position 63) LT-R72 (where arginine is substituted for the wild-type amino acid at position 72), CT-S109 (where serine is substituted for the wild-type amino acid at position 109), and PT-K9/G129 (where lysine is substituted for the wild-type amino acid at position 9 and glycine substituted at position 129) (see, e.g., International Publication Nos. W093/13202 and W092/19265); and (7) other substances that act as immunostimulating agents to enhance the effectiveness of the composition.

Muramyl peptides include, but are not limited to, N-acetyl-muramyl-L-threonyl-D-isoglutamine (thr-MDP), N-acteyl-normuramyl-L-alanyl-D-isogluatme (nor-MDP), N-acetylmuramyl-L-alanyl-D-isogluatminyl-L-alanine-2-(l'-2'-dipalmitoyl-sn-glycero-3-huydroxyphosphoryloxy)-ethylamine (MTP-PE), etc.

Typically, the vaccine compositions are prepared as injectables, either as liquid solutions or suspensions; solid forms suitable for solution in, or suspension in, liquid vehicles prior to injection may also be prepared. The preparation also may be emulsified or encapsulated in liposomes for enhanced adjuvant effect, as discussed above.

The vaccines will comprise a therapeutically effective amount of the modified Env proteins, or complexes of the proteins, or nucleotide sequences encoding the same, and any other of the above-mentioned components, as needed. By "therapeutically effective amount" is meant an amount of a modified Env (e.g., gp120) protein which will induce a protective immunological response in the uninfected, infected or unexposed individual to which it is administered. Such a response will generally result in the development in the subject of a secretory, cellular and/or antibody-mediated immune response to the vaccine. Usually, such

a response includes but is not limited to one or more of the following effects; the production of antibodies from any of the immunological classes, such as immunoglobulins A, D, E, G or M; the proliferation of B and T lymphocytes; the provision of activation, growth and differentiation signals to immunological cells; expansion of helper T cell, suppressor T cell, and/or cytotoxic T cell.

5

10

15

20

25

30

Preferably, the effective amount is sufficient to bring about treatment or prevention of disease symptoms. The exact amount necessary will vary depending on the subject being treated; the age and general condition of the individual to be treated; the capacity of the individual's immune system to synthesize antibodies; the degree of protection desired; the severity of the condition being treated; the particular Env polypeptide selected and its mode of administration, among other factors. An appropriate effective amount can be readily determined by one of skill in the art. A "therapeutically effective amount" will fall in a relatively broad range that can be determined through routine trials.

Once formulated, the nucleic acid vaccines may be accomplished with or without viral vectors, as described above, by injection using either a conventional syringe or a gene gun, such as the Accell® gene delivery system (PowderJect Technologies, Inc., Oxford, England). Delivery of DNA into cells of the epidermis is particularly preferred as this mode of administration provides access to skin-associated lymphoid cells and provides for a transient presence of DNA in the recipient. Both nucleic acids and/or peptides can be injected either subcutaneously, epidermally, intradermally, intramucosally such as nasally, rectally and vaginally, intraperitoneally, intravenously, orally or intramuscularly. Other modes of administration include oral and pulmonary administration, suppositories, needle-less injection, transcutaneous and transdermal applications. Dosage treatment may be a single dose schedule or a multiple dose schedule. Administration of nucleic acids may also be combined with administration of peptides or other substances.

While the invention has been described in conjunction with the preferred specific embodiments thereof, it is to be understood that the foregoing description as well as the examples which follow are intended to illustrate and not limit the scope of the invention. Other aspects, advantages and modifications within the scope of the invention will be apparent to those skilled in the art to which the invention pertains.

## Experimental

5

10

15

20

25

30

Below are examples of specific embodiments for carrying out the present invention. The examples are offered for illustrative purposes only, and are not intended to limit the scope of the present invention in any way.

Efforts have been made to ensure accuracy with respect to numbers used (e.g., amounts, temperatures, etc.), but some experimental error and deviation should, of course, be allowed for.

#### EXAMPLE 1

A.1. Best-Fit and Homology Searches

The crystal structure of HXB-2 gp 120 was downloaded from the Brookhaven database (COMPLEX (HIV ENVELOPE PROTEIN/CD4/FAB) 15-JUN-98 1GC1 TITLE: HIV-1 GP120 CORE COMPLEXED WITH CD4 AND A NEUTRALIZING HUMAN ANTIBODY). Beta strands 3, 2, 21, and 20 of gp 120 form a sheet near the CD4 binding site. Strands  $\beta$ -3 and  $\beta$ -2 are connected by the V1/V2 loop. Strands  $\beta$ -21 and  $\beta$ -20 are connected by another small loop. The H-bonds at the interface between strands  $\beta$ -2 and  $\beta$ -21 are the only connection between domains of the "lower" half of the protein (joining helix alpha 1 to the CD4 binding site). This beta sheet and these loops mask some antigens (e.g., antigens which may generate neutralizing antibodies) that are only exposed during the CD4 binding.

Constructs that remove enough of the beta sheet to expose the antigens in the CD4 binding site, but leave enough of the protein to allow correct folding were designed. Specifically targeted were modifications to the small loop and, optional deletion of the V1/V2 loops. Three different types of constructs were designed: (1) constructs encoding polypeptides that leave the number of residues making up the entire 4-strand beta sheet intact, but replace one or more residues; (2) constructs that encode polypeptide having at least one residue of at least one beta strand excised or (3) constructs encoding polypeptides having at least two residues of at least one beta strand excised. Thus, a total of 6 different turns were needed to rejoin the ends of the strands.

Initially, residues in the small loop (residues 427-430, relative to HXB-2) and connected beta strands ( $\beta$ -20 and  $\beta$ -21) were modified to contain Gly and Pro (common in beta turns). These sequences were then used as the target to match in each search. The

geometry of the target was matched to known proteins in the Brookhaven Protein Data Bank. In particular, 5-residue turns (including an overlapping single residue at the N-terminal, the 2 residue target turn and 2 overlapping residues at the C-terminal) were searched in the databases. In other words, these modified loops add a 2 residue turn that should be able to support a geometry that will maintain the beta-sheet structure of the wild type protein. The calculations were performed using the default parameters in the Loop Search feature of the Biopolymer module of the Syby1 molecular modeling package. In each case, the 25 best fits based on geometry alone were reviewed and, of those, several selected for homology and fit.

In addition, it was also determined what modifications could be made to remove most of the V1/V2 loop (residues 124-198, relative to HXB-2) yet leave the geometry of the protein intact. As with the small loop, constructs were also designed which excised one or more residues from the  $\beta$ -2 strand (residues 119-123 of HXB-2), the  $\beta$ -3 strand (residues 199-201 of HXB-2) or both  $\beta$ -2 and  $\beta$ -3. For these constructs, known loops were searched to match the geometry of a pentamer (including two remaining residues from the N-terminal side, a 2 residue turn and 1 C-terminal residue). For these searches, Gly-Gly was preferred as the insert along with at least one C-terminal substitution.

# A.2. Small Loop Replacements

In one aspect, the native sequence was replaced with residues that expose the CD4 binding site, but leave the overall geometry of the protein relatively unchanged. For the small loop replacements, the target to match was: ASN425-MET426-GLY427-GLY428-GLY431. Results of the search are summarized in Table 1.

Table 1: Search of Small Loop (Asn425 through Gly431)

25

5

10

15

20

Rank	Sequence	RMSD	% Homology	Seq Id No.
Best fit	LYS-ASP-SER-ASN-ASN	0.16689	62.5	27
3	TYR-GLY-LEU-GLY-LEU	0.220308	62.5	28
4	GLU-ARG-GLU-ASP-GLY	0.241754	62.5	29
7	ARG-LYS-GLY-GLY-ASN	0.24881	100	30
12	TRP-THR-GLY-SER-TYR	0.26417	83.33	31

30

Based on these results, constructs encoding Gly-Gly (#7), Gly-Ser (#12) or Gly-Gly-Asn (#7) were recommended.

As V1/V2 and one or more residues of  $\beta$ -2 and  $\beta$ -3 are also optionally deleted in the modified polypeptides of the invention, known loops to match the geometry of the V1/V2 loop were also searched. The V1/V2 loop the target to match was: Lys121-Leu-122-Gly123-Gly124-Ser199. Some notable matches are shown in Table 2:

Table 2: Search of V1/V2 loop (Lys121 through Ser199)

Rank	Sequence	RMSD	% Homology	Seq Id. No.
Best fit	GLN-VAL-HIS-ASP-GLU	0.154764	68.75	32
2	LYS-GLU-GLY-ASP-LYS	0.15718	81.25	33
9	ARG-SER-GLY-ARG-SER	0.173731	68.75	34
11	THR-LEU-GLY-ASN-SER	0.175554	81.25	35
16	HIS-PHE-GLY-ALA-GLY	0.178772	93.75	36

15

10

5

Based on these searches, constructs encoding Gly-Asn in place of V1/V2 were recommended.

### A.3. One Additional Residue Excisions

20

For a slightly truncated small loop, one more residue was trimmed from each beta strand to slightly shorten the beta sheet. The target to match was: ILE424-ASN425-GLY426-GLY427-LYS432. Results are shown in Table 3:

Table 3: Search of Beta sheet shortened by One residue (Ile424 through Lys432)

25

Rank	Sequence	RMSD	% Homology	Seq Id No.
Best fit:	ARG-MET-ALA-PRO-VAL	0.316805	58.33	37
Best	ASP-SER-ASP-GLY-PRO	0.440896	83.33	38
hom:	_	ļ	<u> </u>	

Although these searches showed more variation and worse fits than the previous truncation, the Pro-Val or Pro-Leu encoding constructs were very similar. Accordingly, Ala-Pro encoding constructs were recommended.

Sequences encoding gp120 polypeptides having V1/V2 deleted and an additional residue from  $\beta$ -2 or  $\beta$ -3 excised were also searched. The V1/V2 loop the target to match was: VAL120-LYS121-GLY122-GLY123-VAL200. Some notable matches are shown in Table 4.

Table 4: Search of V1/V2 loop (Val120 through Val200)

10	Rank	Sequence	RMSD	% Homology	Seq Id No
	Best fit:	THR-VAL-ASP-PRO-TYR	0.400892	58.33333	39
	2	SER-THR-ASN-PRO-LEU	0.402575	54.16667	40
	3	THR-ARG-SER-PRO-LEU	0.403965	58.33333	41
	7	ARG-MET-ALA-PRO-VAL	0.440118	58.33333	42

15

20

25

5

The construct encoding Ala-Pro (e.g., #7) was recommended.

# A.4. Further Excisions

In yet another truncation, an additional residue was trimmed from the  $\beta$ -20 and  $\beta$ -21 strands to further shorten the beta sheet. The target to match was ILE423-ILE424-GLY425-GLY426-ALA433. Notable matches are shown in Table 5.

Table 5: Search of Beta sheet shortened by Two Residues (Ile423 through Ala433)

Rank	Sequence	RMSD	% Homology	Seq Id No
Best fit:	THR-TYR-GLU-GLY-VAL	0.130107	79.16666	43
2	GLN-VAL-GLY-ASN-THR	0.138245	79.16666	44
3:	THR-VAL-GLY-GLY-ILE	0.153362	100	45

A construct encoding Gly-Gly (e.g., #3), which has 100% homology, was

30 recommended.

Also searched were sequences encoding a deleted V1/V2 region and at least two residues excised from  $\beta$ -2,  $\beta$ -3 or at least one residue excised from  $\beta$ -2 and  $\beta$ -3. The target to match was: CYS119-VAL120-GLY121-GLY122-ILE201. Notable matches are shown in Table 6.

5

Table 6: Search of V1/V2 loop (Cysl 19 through Ile201)

Rank	Sequence	RMSD	% Homology	Seq Id No
Best fit:	ASP-LEU-PRO-GLY-CYS	0.250501	75	46
4	ASP-VAL-GLY-GLY-LEU	0.290383	100	47

10

15

It was determined that both constructs would be used.

## B.1. Constructs encoding modified Env polypeptides

As described above, the native loops extruding from the 4- $\beta$  antiparallel-stands were excised and replaced with 1 to 3 residue turns. The loops were replaced so as to leave the entire  $\beta$ -strands or excised by trimming one or more amino acid from each side of the connected strands. The ends of the strands were rejoined with turns that preserve the same backbone geometry (e.g., tertiary structure of  $\beta$ -20 and  $\beta$ -21), as determined by searching the Brookhaven Protein Data Bank.

20

Table 7A is a summary of the truncations of the variable regions 1 and 2 recommended for this study, as determined in Example 1.A. above.

Table 7A

V1/V2 Modifications	SEQ ID NO	Figure
-LEU122-GLY-ASN-SER199	7	10
-LYS121-ALA-PRO-VAL200-	6	9
-VAL120-GLY-GLY-ILE201-	4	7
-VAL120- <b>PRO-GLY-</b> ILE201B-	5	8
-VAL120-GLY-ALA-GLY-ALA204-	3	6
-VAL120-GLY-GLY-ALA-THR202-	8	11
-VAL127-GLY-ALA-GLY-ASN195-	25	28

10

15

5

As previously noted, the polypeptides encoded by the constructs of the present invention are numbered relative to HXB-2, but the particular amino acid residue of the polypeptides encoded by these exemplary constructs is based on SF-162. Thus, for example, although amino acid residue 195 in HXB-2 is a serine (S), constructs encoding polypeptides having then wild type SF162 sequence will have an asparagine (N) at this position. Table 7B shows just three of the variations in amino acid sequence between strains HXB-2 and SF162. The entire sequences, including differences in residue and amino acid number, of HXB-2 and SF162 are shown in the alignment of Figure 2 (SEQ ID NOs:1 and 2).

20

Table 7B

HXB-2 amino acid number	HXB-2 Residue	SF162 Residue/amino acid number		
128	Serine (S)	Thr (T)/114		
195	Serine (S)	Asn (N)/188		
426	Met (M)	Arg (R)/411		

25

30

Constructs containing deletions in the  $\beta$ -20 strand,  $\beta$ -21 stand and small loop were also constructed. Shown in Table 8 are constructs encoding truncations in these regions. The constructs in Table 8 are numbered relative to HXB-2 but the unmodified amino acid sequence is based on SF162. Thus, the construct encodes an arginine (Arg) as is found in

PCT/US99/31272 WO 00/39303

SF162 in the amino acid position numbered 426 relative to HXB-2 (See, also, Table 7B). Changes from wildtype (SF162) are shown in bold in Table 8B.

Table 8

5

10

15

**SEQ ID NO** Small Loop/ $\beta$ -20 and  $\beta$ -21 (Modified) Figure -TRP427-GLY-GLY431-12 -ARG426-GLY-GLY-GLY431-10 13 14 -ARG426-GLY-SER-GLY431B-11 -ARG426-GLY-GLY-ASN-LYS432-12 15 -ASN425-ALA-PRO-LYS432-13 16 -ILE424-GLY-GLY-ALA433-14 17 -ILE423-GLY-GLY-MET434-15 18 GLN422-GLY-GLY-TYR435-19 16 17 20 -GLN422-ALA-PRO-TYR435B-

The deletion constructs shown in Tables 7 and 8 for each one of the  $\beta$ -strands and combinations of them are constructed. These deletions will be tested in the Env forms gp120, gp140 and gp160 from different HIV strains like subtype B strains (e.g., SF162, US4, SF2), subtype E strains (e.g., CM235) and subtype C strains (e.g., AF110968 or AF110975).

20 Exemplary constructs for SF162 are shown in the

> Figures and are summarized in Table 9. As noted above in Figure 2 and Table 7B, in the bridging sheet region, the amino acid sequence of SF162 differs from HXB-2 in that the Met426 of HXB-2 is an Arg in SF162. In Table 9, V1/V2 refers to deletions in the V1/V2 region; # bsm refers to a modification in the bridging sheet small loop.

25

Table 9				
Construct	Seq. Id.	Fig.	Modification/Amino acid sequence	
Val120-Ala204	3	6	V1/V2: Val120-Gly-Ala-Gly-Ala204	
Val120-Ile201	4	7	V1/V2: Val120-Gly-Gly-Ile201	
Val120-1le201B	5	8	V1/V2: Val120-Pro-Gly-Ile201	
Lys121-Val200	6	9	V1/V2: Lys121-Ala-Pro-Val200	

		T	able 9
Construct	Seq. Id.	Fig.	Modification/Amino acid sequence
Leu122-Ser199	7	10	V1/V2: Leu122-Gly-Asn-Ser199
Val120-Thr202	8	11	V1/V2: Val120-Gly-Gly-Ala-Thr202
Trp427-Gly431	9	12	bsm: Trp427-Gly-Gly431
Arg426-Gly431	10	13	bsm: Arg426-Gly-Gly-Gly431
Arg426-Gly431B	11	14	bsm: Arg426-Gly-Ser-Gly431
Arg426-Lys432	12	15	bsm: Arg426-Gly-Gly-Asn-Lys432
Asn425-Lys432	13	16	bsm: Asn425-Ala-Pro-Lys432
Ile424-Ala433	14	17	bsm: Ile424-Gly-Gly-Ala433
Ile423-Met434	15	18	bsm: Ile423-Gly-Gly-Met434
Gln422-Tyr435	16	19	bsm: Gln422-Gly-Gly-Tyr435
Val127-Asn195	25	28	bsm: Val127-Gly-Ala-Gly-Asn195
Gln422-Tyr435B	17	20	bsm: Gln422-Ala-Pro-Tyr435
Leu122-Ser199; Arg426-Gly431	18	21	V1/V2/bsm: Leu122-Gly-Asn-Ser199 Arg42 Gly-Gly-Gly431
Leu122-Ser199; Arg426-Lys432	19	22	V1/V2/bsm: Leu122-Gly-Asn-Ser199 Arg42 Gly-Gly-Asn-Lys432
Leu122-Ser199-Trp427- Gly431	20	23	V1/V2/bsm: Leu122-Gly-Asn-Ser199 Trp42 Gly-Gly431
Lys121-Val200- Asn425-Lys432	21	24	V1/V2/bsm: Lys121-Ala-Pro-Val200 Asn42 Ala-Pro-Lys432
Val120-Ile201-Ile424- Ala433	22	25	V1/V2/bsm: Val120-Gly-Gly-Ile201 Ile424- Gly-Gly-Ala433
Val120-Ile201B-Ile424- Ala433	23	26	V1/V2/bsm: Val120-Pro-Gly-Ile201 Ile424- Gly-Gly-Ala43
Val120-Thr202; Ile424- Ala433	24	27	V1/V2/bsm: Val120-Gly-Gly-Ala-Thr202 Ile424-Gly-Gly-Ala433
Val127-Asn195; Arg426-Gly431	25	29	V1/V2/bsm: Val127-Gly-Ala-Gly-Asn195 Arg426-Gly-Gly-Gly431

Combinations of V1/V2 deletions and bridging sheet small loop modifications in addition to those specifically shown in Table 9 are also within the scope of the present invention. Various forms of the different embodiments of the invention, described herein, may be combined.

The first screening will be done after transient expression in COS-7, RD and/or 293 cells. The proteins that are expressed will be analyzed by immunoblot, ELISA, and for binding to mAbs directed to the CD4 binding site and other important epitopes on gp120 to determine integrity of structure. They will also be tested in a CD4 binding assay and, in addition, the binding of neutralizing antibodies, for example using patient sera or mAb 448D (directed to Glu370 and Tyr384, a region of the CD4 binding groove that is not altered by the deletions).

5

10

The immunogenicity of these novel Env glycoproteins will be tested in rodents and primates. The structures will be administered as DNA vaccines or adjuvanted protein vaccines or in combined modalities. The goal of these vaccinations will be to archive broadly reactive neutralizing antibody responses.

## Claims:

What is claimed is:

- 1. A polynucleotide encoding a modified HIV Env polypeptide wherein the
  polypeptide has at least one amino acid deleted or replaced in the region corresponding to
  residues 420 to 436 relative to HXB-2 (SEQ ID NO:1).
- The polynucleotide of claim 1, wherein the region corresponding to residues 124 198 relative to HXB-2 is deleted and at least one amino acid is deleted or replaced in the regions corresponding to the residues 119 to 123 and 199 to 210 relative to HXB-2 (SEQ ID NO:1).
- The polynucleotide of claim 1, wherein at least one amino acid in the region
   corresponding to residues 427 through 429 relative to HXB-2 (SEQ ID NO:1) is deleted or replaced.
  - 4. The polynucleotide of claim 2, wherein at least one amino acid of the in the region corresponding to residues 427 through 429 relative to HXB-2 (SEQ ID NO:1) is deleted or replaced.
    - 5. The polynucleotide of claim 1, wherein the amino acid sequence of the modified HIV Env polypeptide is based on strain SF162.
- 25 6. An immunogenic modified HIV Env polypeptide having at least one amino acid deleted or replaced in the region corresponding to residues 420 through 436, relative to HXB-2 (SEQ ID NO:1).
- 7. The polypeptide of claim 6, wherein one amino acid is deleted in the region corresponding to residues 420 through 436, relative to HXB-2 (SEQ ID NO:1).

8. The polypeptide of claim 6, wherein more than one amino acid is deleted in the region corresponding to residues 420 through 436, relative to HXB-2 (SEQ ID NO:1).

- 9. The polypeptide of claim 6, wherein at least one amino acid is replaced in the region corresponding to residues 420 through 436, relative to HXB-2 (SEQ ID NO:1).
  - 10. The polypeptide of claim 6, wherein at least one amino acid residue between about amino acid residue 427 and amino acid residue 429 relative to HXB-2 (SEQ ID NO:1) is deleted or replaced.

10

- 11. The polypeptide of claim 6, wherein the V1 and V2 regions of the polypeptide are truncated.
- 12. The polypeptide of claim 10, wherein the V1 and V2 regions of the polypeptide are truncated.
  - 13. The polypeptide of claim 6, wherein the amino acid sequence of the modified HIV Env polypeptide is based on strain SF162.
- 20 14. A construct comprising the nucleotide sequence depicted in Figure 6 (SEQ ID NO:3).
  - 15. A construct comprising the nucleotide sequence depicted in Figure 7 (SEQ ID NO:4).

- 16. A construct comprising the nucleotide sequence depicted in Figure 8 (SEQ ID NO:5).
- 17. A construct comprising the nucleotide sequence depicted in Figure 9 (SEQ ID30 NO:6).

18. A construct comprising the nucleotide sequence depicted in Figure 10 (SEQ ID NO:7).

- 19. A construct comprising the nucleotide sequence depicted in Figure 11 (SEQ ID5 NO:8).
  - 20. A construct comprising the nucleotide sequence depicted in Figure 12 (SEQ ID NO:9).
- 10 21. A construct comprising the nucleotide sequence depicted in Figure 13 (SEQ ID NO:10).
  - 22. A construct comprising the nucleotide sequence depicted in Figure 14 (SEQ ID NO:11).

23. A construct comprising the nucleotide sequence depicted in Figure 15 (SEQ ID NO:12).

15

- 24. A construct comprising the nucleotide sequence depicted in Figure 16 (SEQ IDNO:13).
  - 25. A construct comprising the nucleotide sequence depicted in Figure 17 (SEQ ID NO:14).
- 25 26. A construct comprising the nucleotide sequence depicted in Figure 18 (SEQ ID NO:15).
  - 27. A construct comprising the nucleotide sequence depicted in Figure 19 (SEQ ID NO:16).
  - 28. A construct comprising the nucleotide sequence depicted in Figure 20 (SEQ ID NO:17).

29. A construct comprising the nucleotide sequence depicted in Figure 21 (SEQ ID NO:18).

- 30. A construct comprising the nucleotide sequence depicted in Figure 22 (SEQ IDNO:19).
  - 31. A construct comprising the nucleotide sequence depicted in Figure 23 (SEQ ID NO:20).
- 32. A construct comprising the nucleotide sequence depicted in Figure 24 (SEQ ID NO:21).
  - 33. A construct comprising the nucleotide sequence depicted in Figure 25 (SEQ ID NO:22).

34. A construct comprising the nucleotide sequence depicted in Figure 26 (SEQ ID NO:23).

35. A construct comprising the nucleotide sequence depicted in Figure 27 (SEQ ID NO:24).

15

- 36. A construct comprising the nucleotide sequence depicted in Figure 28 (SEQ ID NO:25).
- 25 37. A construct comprising the nucleotide sequence depicted in Figure 29 (SEQ ID NO:26).
  - 38. A vaccine composition comprising a polynucleotide encoding a modified Env polypeptide according to any one of claims 1-5.
  - 39. A vaccine composition comprising a polynucleotide construct encoding a modified Env polypeptide according to any of claims 14-37.

40. A vaccine composition comprising a modified Env polypeptide according to any of claims 6-13.

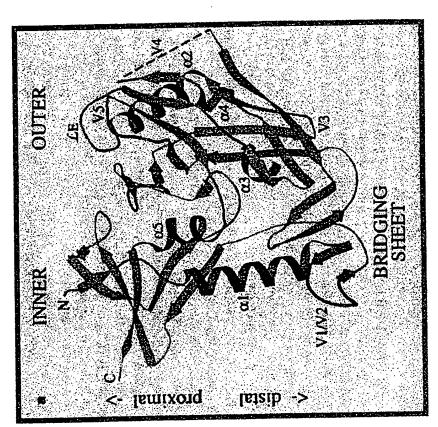
41. The vaccine composition of any of claims 38-40, further comprising an adjuvant.

5

- 42. A method of inducing an immune response in subject comprising, administering a polynucleotide according to any one of claims 1-5 in an amount sufficient to induce an immune response in the subject.
- 43. A method of inducing an immune response in subject comprising, administering a polynucleotide construct according to any one of claims 14-37 in an amount sufficient to induce an immune response in the subject.
- 44. A method of inducing an immune response in a subject comprising administering a composition comprising a modified Env polypeptide according to any one of claims 6-13, wherein the composition is administered in an amount sufficient to induce an immune response in the subject
- 45. The method of any of claims 42-44 further comprising administering an adjuvant to the subject.
  - 46. A method of inducing an immune response in a subject comprising
  - (a) administering a first composition comprising a polynucleotide according to any of claims 1-5 in a priming step and
  - (b) administering a second composition comprising a modified Env polypeptide according to any of claims 6-13, as a booster, in an amount sufficient to induce an immune response in the subject.
- 47. The method of claim 46 wherein the first composition or second composition further comprise an adjuvant.

48. The method of claim 46 wherein the first and second compositions further comprise an adjuvant.

## gp120 core structure



F1G. 1

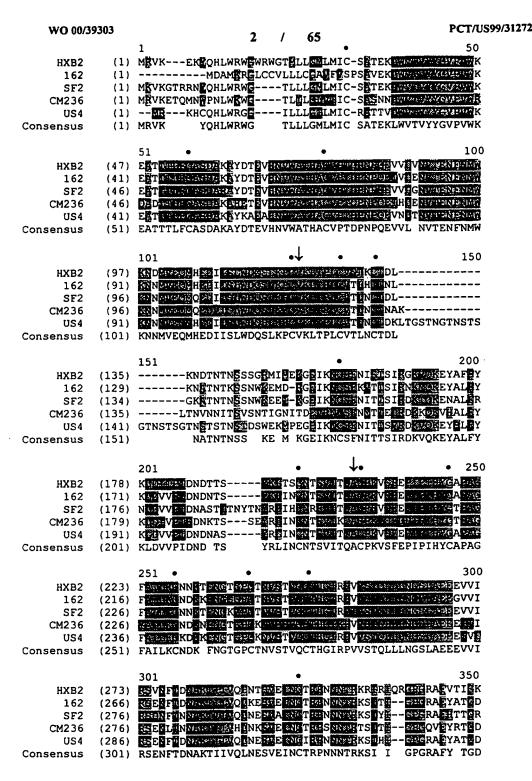


FIG. 2A

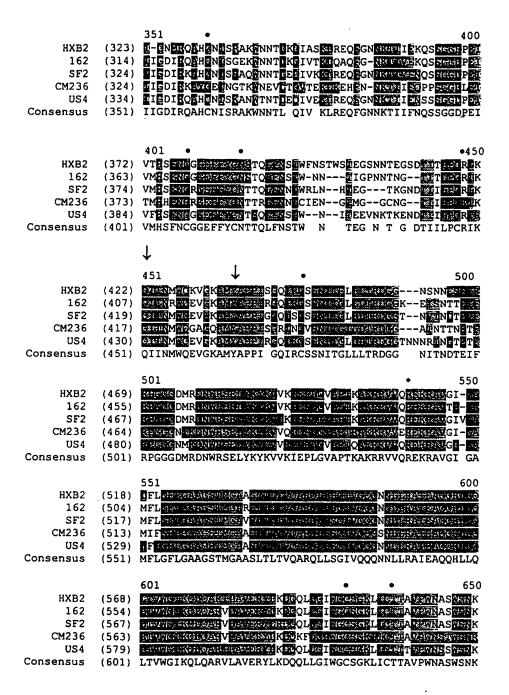


FIG. 2B

FIG. 2C

WO 00/39303			5	1	65		PCT/US99/31272
		1					40
Leu122-Ser199	(1)				TGCAATGAAGA		
Val127-Asn195	(1)				TGCAATGAAGA		
Val120-Ile201B	(1)	GAATTCGC	CACC	ATGG#	TGCAATGAAGA	<u>Ģ</u> ĀGGGCTÇŢĞ	ÇT
Val120-Ala204	(1)				TGCAATGAAGA		
Vall20-Ile201	(1)	GAATTCGC	CACC	ATGGA	TGCAATGAAGA	GAGGGCTCTG	CT
Val120-Thr202	(1)				TGCAATGAAGA		
Lys121-Va1200	(1)	GAATTCGC	CACC	ATGGP	ŢĠĊAATĞAAGA	SAGGGCTCTG	CT
Consensus	(1)	GAATTCGC	CACC	ATGGA	TGCAATGAAGA	GAGGGCTCTG	CT
		41				•	80
Leu122-Ser199	(41)				GGAGCAGTCTT		
Val127-Asn195	(41)				GGAGCAGTCTTC		
Val120-Ile201B	(41)	GTGTGGTG	CTGC	TGTGT	GGAGCAGTCTT	CGTTTCGCCC	AG
Val120-Ala204	(41)	GTGTGCTG	CTGC	TGTGT	GGAGCAGICIT	CGTTTCGCCC	<u>A</u> G
Val120-Ile201	(41)	GTGTGCTG	CTGÇ	TGTGT	GGAGCAGTCTTC	CGTTTCGCCC	AG
Val120-Thr202	(41)	GTGTGCTG	ÇŤĠĊ	TGTGT	GGAGCAGTCTTC	CGTTTCGCCC	ÀG
Lys121-Val200	(41)	GTGTGCTG	CTGC	TGTGT	GGAGCAGTCTTC	CGTTTCGCCC	AG
Consensus	(41)	GTGTGCTG	CTGC	TGTGT	GGAGCAGTCTTC	CGTTTCGCCC	AG
		81				1	20
Leu122-Ser199	(81)				<b>GGGTGACCGTGT</b>		
Val127-Asn195	(81)	ceccefee	AGAA	GCTGT	GGGTGACCGTG1	PACTACGGCG	IG
Val120-Ile201B	(81)	CCCCCTGG	<u>A</u> GAA	CCIGI	GGGTGACCGTGT	PACTACGGCG	rG.
Val120-Ala204	(81)	CGCCGTGG	AGAA	GCIGI	<b>GEGTGACCGTGT</b>	TACTACGGCG	rG
Val120-Ile201	(81)	CCCCGTGG	AGAA	GCIGT	<b>GGGTGACCGTG</b> I	ACTACGGCG	rg
Val120-Thr202	(81)	CGCCGTGG	ÂĞAA	GCTGT	<b>GGGTGACCGTGT</b>	ACTACGGCG	rg
Lys121-Va1200	(81)	CCCCCTGG	AGAA	GCTGT	GGGTGACCGTGT	ACTACGGCG	ŢĢ
Consensus	(81)	CGCCGTGG	AGAA	GCTGT	GGGTGACCGTGT	'ACTACGGCG'	rg
		121					60
Leu122-Ser199	(121)				CÁCCACCACCCI		
Val127-Asn195	(121)				CACCACCACCCI		
Vall20-Ile201B	(121)				CACCACCACCCI		
Val120-Ala204	(121)	CCCGTGTG	JAAG	GAGGC	CACCACCACCCI	GTTCTGCGC	DA .
Val120-Ile201	(121)				CACCACCACCCI		
Val120-Thr202	(121)	elete steriet	AAG	ch (ccc	CACCACCACCCT	GTTCTGCGC	ZA .
Lys121-Val200	(121)	coccusic	AAG	GAGGE	CACCACCCT	GTICTGCGC	ĈĀ.
Consensus	(121)	CCCGTGTG	GAAG	GAGGC	CACCACCACCCT		
		161					00
Leu122-Ser199	(161)				GACACCGAGGTG		
Val127-Asn195	(161)				GACACCGAGGTG		
Val120-Ile201B	(161)				SACACCGAGGTG		
Val120-Ala204	(161)				CACCEAGGIG		
Val120-Ile201	(161)				CACCCAGGTG		
Val120-Thr202	(161)	GEGACGUL	VAGG	CTAC	sacaccgaegte	CACAACGTG	i.C
Lys121-Val200	(161)				<b>TACACCGAGGTG</b>		
Consensus	(161)	GCGACGCCA	AGGG	CCTAC	GACACCGAGGTG	CACAACGTG	rg
		201				24	-
Leul22-Ser199	(201)	GGCCACCG?	(CCC)	TIGCG:	TGECCACCGACC	CCAACCCCC	<sup>j</sup> G
Val127-Asn195	(201)	GGCCACCCA	CGC	TGCG	r <b>GCCCACCGACC</b>	CCAACCCCC	\G
Val120-Ile201B	(201)	GCCCACCCA	CCCC	TGCG	IGCCCACCGACC	CCAACCCCC	<b>ķ</b> G
Val120-Ala204	(201)	GGCCACGC?	CGC	TGCG	(GCCCACCGACC	CCAACCCCC	\G
Val120-Ile201	(201)	GGCCACGCA	CGCC	TGÇG	TGCCCACCGACC	CCAACCCCC	iœ
Val120-Thr202	(201)	GCCACCCA	CGCC	TGCG:	FGCCCACCGACC	CCAACCCCC	<b>V</b> G
Lys121-Va1200	(201)	GGCCACCCA	CCCC	;TGCG;	GCCCACCGACC	CCAACCCCC	<i>I</i> G
Consensus	(201)		CGCC	CTGCG	TGCCCACCGACC		
		241				. 28	
Leul22-Ser199					GTGACCGAGAA		
Val127-Asn195	(241)	GAGATCGTG	CTGC	AGAAC	CGTGACCGAGAA	CTTCAACATO	ST

WO 00/39303		6	1	65	P	CT/US99/31272
Val120-Ile201B	(241)	GAGATCGTGCTG	AGAAC	GTGACCGAG	ÄACTTCÄACATG	T
Val120 1122012		GAGATCGTGCTGG				
Val120-Ile201	(241)	GAGATCGTGCTG	AGAAC	GTGACCGAG	AACTTCAACATG	T
Val120-Thr202		GAGATCGTGCTGC				
Lys121-Val200		GAGATCGTGCTGG				
Consensus		GAGATCGTGCTG				
Consensus	(2.11)	281			32	
Leu122-Ser199	(281)	GGAAGAACAACAT	GGTGG	AGCAGATGC		
Val127-Asn195		GGAAGAACAACAT				
Val120-Ile201B		GGAAGAACAACAT				
Val120-Ala204		GGAAGAACAACAT				
Val120-Ile201		GGAAGAACAACAT				
Val120 Thc201		GGAAGAACAACAT				
Lys121-Val200	(281)					
Consensus		GGAAGAACAACAT	GGTGG	AGCAGATGC	ACGAGGACATCA	T
Consensus	(201)	321			36	
Leul22-Ser199	(321)	CAGCCTGTGGGAC	CAGAG	CCTGAAGCC		_
Val127-Asn195	(321)	CAGCCTGTGGGAC	CAGAG	CCTGAAGCC	CTGCGTGAAGCT	Ğ
Val120-Ile201B	(321)	CAGCCTGTGGGAC	CAGAG	CCTGAAGCC	CTGCGTGCC	-
Val120-1162016 Val120-Ala204		CAGCCTGTGGGAC				
Val120-A1a204 Val120-Ile201	(321)	CAGCCTGTGGGAC	CACAG	CCTGAAGCC	CTECETEGE	<b>-</b>
Val120-The201 Val120-Thr202	13211	CAGCCTGTGGGAC	CACAC	CCTGAAGCC	CTGCGTGGG	-
Lys121-Val200	(321)	CAGCCTGTGGGAC	CAGAG	CCTGAAGCC	CTGCGTGAAGG-	_
Consensus	(321)	CAGCCTGTGGGAC	CACAC	CCTGAAGCC	CTGCGTG	
Consensus	(321)	361	C110110		40	0
Leu122-Ser199	(361)			GCAA		
Val127-Asn195	(361)	ACCCCCTGTGCG	TGGGG	GCAGGGAAC'	TGCAACACCAGC	Ğ
Val120-Ile201B	(357)				C	G
Val120-Ala204	(357)					_
Val120 Ala204 Val120-Ile201	(357)				с	G
Val120 Thc201 Val120-Thr202	(357)				<u>-</u>	G
Lys121-Val200	(359)				ccccc	G
Consensus	(361)				С	
co	(552)	401			44	0
Leu122-Ser199	(371)	TGATCACCCAGGC	CTGCC	CCAAGGTGA	GCTTCGAGCCCA	Ť
Val127-Asn195	(401)	and the second of the second of the second				
Val120-Ile201B	(359)	The second secon				
Val120-Ala204	(357)					
Val120-Ile201		GCATCACCCAGGC				
Val120 Thr202		GCGCCACCCAGGC				
Lys121-Val200	(365)	TGATCACGCAGGC				
Consensus	(401)	ATCACCCAGGC	CTGCC	CCAAGGTGA	GCTTCGAGCCCA	T
0000	( /	441			48	
Leu122-Ser199	(411)	CCCCATCCACTAC	recec	CECEGEGG	CTTCGCCATCCT	G
Val127-Asn195	(441)	CCCCATCCACTAG	TECEC	GECCECCGG	TTCGCCATCCT	Ğ
Val120-Ile201B	(399)	CCCCATCCACTAC	TGCGC	CCCCGGCGG	TTCGCCATCGT	Ğ
Val120-Ala204	(393)					
Val120-Ile201	(399)	CCCCATGCACTAC	TGCGC	CCCCCCGG	TTCGCCATCCT	G
Val120-Thr202	(399)	CCCCATCCACTAC	TGCGC	CCCCCCCG	CTTCGCCATCCT	G
Lys121-Val200	(405)					
Consensus	(441)	CCCCATCCACTAC	rgcgc	CCCGCCGG	CTTCGCCATCCT	G
22.72277243		481			52	0
Leu122-Ser199	(451)	AAGTGCAACGACA	AGAAG'	TTCAACGGC	AGCGGCCCCTGC	A
Val127-Asn195	(481)	AAGTGCAACGACA	AGAAG'	TTCAACGGC	AGCGGCCCCTGC	A
Val120-Ile201B	(439)	AAGTGCAACGACA	AGAAG'	TTCAACGGC	AGCGGCCCCTGC	A
Val120 11e2018 Val120-Ala204	(433)	AAGTGCAACGACA	AGAAG'	TTCAACGGC	AGCGGCCCCTGC	A
Val120-A1a204 Val120-Ile201	(439)		AGAAG'	TTCAACGGC	AGCGGCCCCTGC	A
	,			in seem teen to		

(691) ATCACCATCGGCCCGGCCGCCTTCTACGCCACCGGCG Leu122-Ser199 (721) ATCACCATCGGCCCGGCCGCGCCTTCTACGCCACCGGCG Val127-Asn195 (679) ATCACCATCGGCCCCGGCCGCCCTTCTACGCCACCGGCG Val120-Ile201B (673) ATCACCATCGGCCCGGCCGCCCTTCTACGCCACCGGCG Val120-Ala204 (679) ATCACCATCGGCCCCGGCCGCCTTCTACGCCACCGGCG Val120-Ile201 (679) ATCACCATCGGCCCGGCCGCGCTTCTACGCCACCGGCG Val120-Thr202 (685) ATCACCATCGGCCCGGCCGCCCTTCTACGCCACCGGCG Lys121-Val200 (721) ATCACCATCGGCCCGGCCGCCCTTCTACGCCACCGGCG Consensus

Val127-Asn195

Val120-Ala204

Val120-Ile201

Val120-Thr202

Lys121-Val200

Consensus

(633) (639)

(639)

Val120-Ile201B

(681) GATCAACTGCACCCGCCCCAACAACAACACCCCGCAAGAGC

(639) GATCAACTGCACCCGCCCCAACAACAACACCCGCAAGAGC

(645) GATCAACTGCACCCGCCCCAACAACAACACCCCGCAAGAGC

(681) GATCAACTGCACCCGCCCCAACAACAACACCCCGCAAGAGC

GATCAACTGCACCGGCCCGAAGACAACAACACCCGCAAGAGC

```
(959) ACAACACCAACGGCACCATCACCCTGCCCTGCCGCATCAA
Val120-Ile201E
 Val120-Ala204
                 (953) ACAACACCAACGGCACCATCACCCTGCCCTGCCGCATCAA
                 (959) ACAACACCAACGGCACCATCACCCTGCCCTGCCGCATCAA
 Val120-Ile20i
                 (959) ACAACACCAACGCACCATCACCCTGCCCTGCCGCATCAA
 Val120-Thr202
                 (965) ACAACACCAACGGCACCATCACCCTGCCCTGCCGCATCAA
 Lys121-Val200
                (1001) ACAACACCAACGGCACCATCACCCTGCCCTGCCGCATCAA
     Consensus
                (1011) GCAGATCATCAACCGCTGGCAGGAGGTGGGCAAGGCCATG
 Leu122-Ser199
                (1041) GCAGATCATCAACCGCTGGCAGGAGGTGGGCAAGGCCATG
 Val127-Asn195
                 (999) GCAGATCATCAACCGCTGGCAGGAGGTGGGCAAGGCCATG
Val120-Ile201B
                 (993) GCAGATCATCAACCGCTGGCAGGAGGTGGGCAAGGCCATG
 Val120-Ala204
                 (999) GCAGATCATCAACCGCTGGCAGGAGGTGGGCAAGGCCATG
 Val120-Ile201
                 (999) GCAGATCATCAACCGCTGGCAGGAGGTGGGCAAGGCCATG
 Val120-Thr202
                (1005) GCAGATCATCAACCGCTGGCAGGAGGTGGGCAAGGCCATG
 Lys121-Val200
                (1041) GCAGATCATCAACCGCTGGCAGGAGGTGGGCAAGGCCATG
     Consensus
                       1081
               (1051) TACGCCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGCA
 Leu122-Ser199
               (1081) TACGCCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGCA
 Val127-Asn195
                (1039) TACGCCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGCA
Val120-Ile201B
               (1033) TACGCCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGCA
 Val120-Ala204
               (1039) TACGCCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGCA
 Val120-Ile201
               (1039) TACGCCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGCA
 Val120-Thr202
               (1045) TACGCCCCCCCATCCGCGCCAGATCCGCTGCAGCAGCA
 Lys121-Val200
                (1081) TACGCCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGCA
     Consensus
                       1121
                (1091) ACATCACCGGCCTGCTGCTGACCCGCGACGGCGGCAAGGA
 Leu122-Ser199
                (1121) ACATCACCGGCCTGCTGCTGACCCGCGACGGCGGCAAGGA
 Val127-Asn195
                (1079) ACATCACCGGCCTGCTGCTGACCCGCGACGGCGCAAGGA
Val120-Ile201B
                (1073) ACATCACCGGCCTGCTGCTGACCCGCGACGGCGGCAAGGA
 Val120-Ala204
                (1079) ACATCACCGGCCTGCTGCTGACCCGCGACGGCGCAAGGA
 Val120-Ile201
               (1079) ACATCACCGGCCTGCTGACCCGCGACGGCGCAAGGA
 Val120-Thr202
               (1085) ACATCACCGGCCTGCTGCTGACCCGCGACGGCGGCAAGGA
 Lys121-Val200
               (1121) ACATCACCGGCCTGCTGCTGACCCGCGACGGCGGCAAGGA
     Consensus
                                                          1200
               (1131) GATCAGCAACACCACCGAGATCTTCCGCCCCGGCGGCGGC
 Leu122-Ser199
               (1161) GATCAGCAACACCACCGAGATCTTCCGCCCCGGCGGCGGC
 Val127-Asn195
               Val120-Ile201B
 Val120-Ala204
                      GATCAGCAACACCACCGAGATCTTCCGCCCCGGCGGCGC
 Val120-Ile201
               (1119)
               (1119) GATCAGCAACACCACCGAGATCTTCCGCCCCGGCGGCGGC
 Val120-Thr202
               (1125) GATCAGCAACACCACCGAGATCTTCCGCCCCGGCGGCGGC
 Lys121-Val200
              (1161) GATCAGCAACACCACCGAGATCTTCCGCCCCGGCGGCGGC
     Consensus
                      1201
                      GACATGCGCGACAACTGGCGCGCGAGCTGTACAAGTACA
 Leul22-Ser199
               (1171)
                      GACATGCGCGACAACTGGCGCGCGGGCTGTACAAGTACA
               (1201)
 Val127-Asn195
                      GACATGCGCGACAACTGGCGCAGCGAGCTGTACAAGTACA
Val120-Ile201B
               (1159)
               (1153) GACATGCGCGACAACTGGCGCAGCGAGCTGTACAAGTACA
 Val120-Ala204
               (1159) GACATGCGCGACAACTGGCGCGCGAGCTGTACAAGTACA
 Val120-Ile201
                      GACATGCGCGACAACTGGCGCAGCGAGCTGTACAAGTACA
 Val120-Thr202
               (1159)
               (1165) GACATGCGCGACAACTGGCGCGCGGGCTGTACAAGTACA
Lys121-Va1200
    Consensus (1201) GACATGCGCGACAACTGGCGCAGCGAGCTGTACAAGTACA
                                                          1280
                      1241
               (1211) AGGTGGTGAAGATCGAGCCCCTGGGCGTGGCCCCCACCAA
Leu122-Ser199
               (1241) AGGTGGTGAAGATCGAGCCCCTGGGCGTGGCCCCCACCAA
Val127-Asn195
               (1199) AGGTGGTGAAGATCGAGCCCCTGGGCGTGGCCCCCACCAA
Val120-Ile201B
               (1193) AGGTGGTGAAGATCGAGCCCCTGGGCGTGGCCCCCACCAA
Val120-Ala204
               (1199) AGGTGGTGAAGATCGAGCCCCTGGGCGTGGCCCCCACCAA
Val120-Ile201
```

```
1521
                 (1491) GCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTG
 Leul22-Ser199
                 (1521) GCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTG
 Val127-Asn195
                 (1479) GCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTG
Val120-Ile201B
                 (1473) GCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTG
 Val120-Ala204
                 (1479) GCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTG
 Val120-Ile201
 Val120-Thr202
                 (1479) GCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTG
                 (1485) GCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTG
 Lys121-Val200
     Consensus
                 (1521) GCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTG
                         1561
                                                              1600
                 (1531) GGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCG
 Leu122-Ser199
                 (1561) GCCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCG
(1519) GCCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCG
 Val127-Asn195
Val120-Ile201B
                 (1513) GGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCG
 Val120-Ala204
                 (1519) GGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCG
 Val120-Ile201
                 (1519) GGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCG
 Val120-Thr202
                (1525) GGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCG
 Lvs121-Val200
                (1561) GGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCG
     Consensus
                                                              1640
                 (1571) CCGTGCCTGGAACGCCAGCTGGAGCAACAAGAGCCTGGA
 Leu122-Ser199
                 (1601) CCGTGCCCTGGAACGCCAGCTGGAGCAACAAGAGCCTGGA
 Val127-Asn195
                 (1559) CCGTGCCCTGGAACGCCAGCTGGACCAACAAGAGCCTGGA
Val120-Ile201B
                 (1553) CCGTGCCTGGAACGCCAGCTGGAGCAACAAGAGCCTGGA
 Val120-Ala204
                 (1559) CCGTGCCCTGGAACGCCAGCTGGAGCAACAAGAGCCTGGA
 Val120-Ile201
                 (1559) CCGTGCCCTGGAACGCCAGCTGGAGCAACAAGAGCCTGGA
 Val120-Thr202
                (1565) CCGTGCCCTGGAACGCCAGCTGGAGCAACAAGAGCCTGGA
 Lys121-Val200
                 (1601) CCGTGCCCTGGAACGCCAGCTGGAGCAACAAGAGCCTGGA
     Consensus
                        1641
                                                              1680
                 (1611) CCAGATCTGGAACAACATGACCTGGATGGAGTGGGAGCGC
 Leu122-Ser199
 Val127-Asn195
                 (1641) CCAGATCTGGAACAACATGACCTGGATGGAGTGGGAGCGC
                 (1599) CCAGATCTGGAACAACATGACCTGGATGGAGTGGGAGCGC (1593) CCAGATCTGGAACAACATGACCTGGATGGGAGTGGGAGCGC
Val120-Ile201B
 Val120-Ala204
                 (1599) CCAGATCTGGAAGAACATGACCTGGATGGAGTGGGAGGGC
Val120-Ile201
                 (1599) CCAGATCTGGAACAACATGACCTGGATGGAGTGGGAGCGC
 Val120-Thr202
                (1605) CCAGATCTGGAACAACATGACCTGGATGGAGTGGGAGCGC
Lys121-Val200
                (1641) CCAGATCTGGAACAACATGACCTGGATGGAGTGGGAGCGC
     Consensus
                (1651) GAGATCGACAACTACACCAACCTGATCTACACCCTGATCG
Leu122-Ser199
Val127-Asn195
                (1681) GAGATEGACAACTACACCAACCTGATCTACACCCTGATCG
                (1639) GAGATCGACAACTACACCAACCTGATCTACACCCTGATCG
Val120-Ile201B
                (1633) GAGATEGACAACTACAACGAACGTGATCTACACCCTGATCG
Val120-Ala204
                        GAGATUGAGAACTACACCAACCTGATCTACACCCTGATCG
Val120-Ile201
                        GAGATICGACAACTACACCAAGCTCATETACACCCTGATCG
Val120-Thr202
                (1639)
                        GAGATCGACAACTACACCAACCTGATCTACACCCTGATCG
Lys121-Val200
                (1645)
                (1681) GAGATCGACAACTACACCAACCTGATCTACACCCTGATCG
     Consensus
                (1691) AGGAGAGCCAGAACCAGCAGGAGAACGAGCAGGAGCT
Leu122-Ser199
                (1721) AGGAGAGCCAGAACCAGCAGGAGAACAACGAGCAGGAGCT
Val127-Asn195
                (1679) AGGAGAGCCAGAACCAGCAGGAGAAGAACGAGCAGGAGCT
Val120-Ile201B
                (1673) AGGAGACCAGAACCAGCAGGAGAACGAGCAGGAGCT
Val120-Ala204
                (1679) AGGAGAGCCAGAACCAGCAGGAGAAGAACGAGCAGGAGCT
Val120-Ile201
                (1679) AGGAGAGCCAGAACCAGCAGGAGAACGAGCAGGAGCT
Val120-Thr202
                (1685) AGGAGAGCCAGAACCAGCAGGAGAAGAACGAGCAGGAGCT
Lvs121-Val200
     Consensus
                (1721) AGGAGAGCCAGAACCAGCAGGAGAACGAGCAGGAGCT
Leu122-Ser199
                (1731) GCTGGAGCTGGACAAGTGGGCCAGCCTGTGGAACTGGTTC
                (1761) GCTGGAGCTGGACAAGTGGGCCAGCCTGTGGAACTGGTTC
```

```
Val120-Thr202
                                                                               (1959) CGACCGCGACCGCAGCAGCCCCCTGGTGCACGGCCTGCTG
          Lys121-Val200
                                                                                (1965) CGACCGCGACCGCAGCAGCCCCCTGGTGCACGGCCTGCTG
                               Consensus
                                                                                    (2001) CGACCGCGACCGCAGCAGCCCCCTGGTGCACGGCCTGCTG
                                                                                 (2011) GCCCTGATCTGGGACGACCTGCGGAGCCTGTGCCTGTTCA
         Leu122-Ser199
                                                                                 (2041) GCCCTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTCA
          Val127-Asn195
                                                                                 (1999) GCCCTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTCA
     Val120-Ile201B
         Val120-Ala204
                                                                                  (1993) GCCCTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTCA
                                                                                  (1999) GCCCTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTCA
(1999) GCCCTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTCA
         Val120-Ile201
         Val120-Thr202
                                                                                   (2005) GCCCTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTCA
         Lys121-Val200
                                                                                   (2041) GCCCTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTCA
                              Consensus
                                                                                                                       2081
                                                                                   (2051) GCTACCACCGCCTGCGCGACCTGATCCTGATCGCCGCCCG
         Leu122-Ser199
                                                                                 (2081) GCTACCACCECCTGCGCGACCTGATCCTGATCGCCGCCCG
        Val127-Asn195
                                                                             Val120-Ile201B
        Val120-Ala204
                                                                                   (2039) GCTACCACCGCCTGCGCGAGCTGATCCTGATCGCCGCCGG
       Val120-Ile201
        Val120-Thr202
                                                                                 (2039) GETAGRAGEGECTECGCGAGETGATGCTGATCGCCGCCCG
(2045) GCTAGCACCGCCTGCGCGAGCTGATCGTCATCGCCGCCCG
       Lys121-Val200
                                                                                 (2081) GCTACCACCGCCTGCGCGACCTGATCCTGATCGCCGCCCG
                            Consensus
                                                                                 (2091) CATEGYSGASCTGCTGGGCCGCGGGGGCTGGGAGGCCCTG
(2121) CATEGYGGACCTGCTGGGAGGCCCTG
       Leu122-Ser199
       Val127-Asn195
                                                                               (2079) CATUSTEGASCIECTES SCUSCOS SESSCIEGEAGGCCCTG
(2073) CATCETGASCIECTEGECEGCEGCTGGCAGGCCCTG
(2079) CATCETGASCIECTEGECEGCEGCTGGGAGGCCCTG
  Val120-Ile201B
                                                                            (2079)
       Val120-Ala204
      Val120-Ile201
                                                                                (2079) CATCGTGGAGCTGGTGGGCCGCGGGGGGGGCCCTG (2085) CATCGTGGAGCTGGTGGGCCGCGCGGGGGGGGGGCCCTG
       Val120-Thr202
      Lys121-Val200
                                                                               (2121) CATCGTGGAGCTGCTGGGCCGCCGCGGCTGGGAGGCCCTG
                          Consensus
                                                                                (2131) ANGUAGUSE CONTROL GO TO TROL GENERAL CAGGAGO (2161) ANGUAC TEGESTATOC, GC TEGACUS CAGGAGO CONTROL GO TEGACUS CAGGAGO CONTROL CAGGAGO CAGGAGO CONTROL CAGGAGO CAGGAGO CONTROL CAGGAGO CO
     Leu122-Ser199
      Val127-Asn195
 Val120-Ile201B
                                                                              (2119)
                                                                                                                   AAchv.\cuvecccc.p.ccc.ec.ec.ec.crv.covcca.precevecca.precevecca.
                                                                             (2113) AGUAFTEGGGGAAGOTEGUIGEAGTACHIGGATECAGGAGC
     Val120-Ala204
     Val120-Ile201
                                                                            (2119) Avactor of the control of the
                                                                            (2119) MAGNACTGGGGCANGSTGCTGCAGTACTGGATCCAGGAGC
(2125) MAGNACTGGGCANACTGCTGCTGCAGTACTGGATCCAGGAGC
     Val120-Thr202
     Lys121-Val200
                        Consensus
                                                                            (2161) AAGTACTGGGGCAACCTGCTGCAGTACTGGATCCAGGAGC
                                                                                                                   2201
    Leu122-Ser199
                                                                               (2171)
                                                                                                                 THATA WATERS OF REVERSE HEAR OF THE SECURITY O
    Val127-Asn195
                                                                               (2201)
                                                                                                                 HEIMAGNATORACCESCUCERCE CONCERNATION COACCESCATION CORRESPONDENCE OF THE COMPANY 
 Val120-Ile201B
                                                                               (2159)
                                                                                                                 TGD 20 TO THE WORLD HE WILL BE THE CONTROL OF THE PROPERTY OF 
    Val120-Ala204
                                                                               (2153)
                                                                            (2159) PGNAGNACAGGGCGCPTGAGGGSVCFFFGGGGCGATCGCCAT
     Val120-Ile201
    Val120-Thr202
                                                                            (2165) TGAAGAACAGCGCGTGAGCCTGTTCGACGCCATCGCCAT
    Lys121-Val200
                                                                                                              TGAAGAACAGCGCCGTGAGCCTGTTCGACGCCATCGCCAT
                        Consensus
                                                                              (2201)
                                                                            (2211) CGCCGTGGCCGAGGGCACCGACCGCATCATCGAGGTGGCC
  Leul22-Ser199
   Val127-Asn195
                                                                            (2241) CGCCTGGCCGAGGGCACCGACCGCATCATCGAGGTGGCC
                                                                            (2199) CGCCETGGCCGAGGGCACCGACCGCATCATCGAGGTGGCC (2193) CGCCGTGGCCGAGGGCACCGACCGCATCATCGAGGTGGCC (2199) CGCCGTGGCCGAGGGCACCGACCGCATCATCGAGGTGGCC
Val120-Ile201B
   Val120-Ala204
   Val120-Ile201
                                                                             (2199) CGCCGTGGCCGAGGGGACCGATCATCGAGGTGGCC
  Val120-Thr202
  Lys121-Val200
                                                                             (2205) CGCCGTGGCCGAGGGCACCGACCGCATCATCGAGGTGGCC
                                                                        (2241) CGCCGTGGCCGAGGGCACCGACCGCATCATCGAGGTGGCC
                       Consensus
```

```
2281
 Leu122-Ser199 (2251) CAGCGCATCGGCCGCGCCTTCCTGCACATCCCCGCCGCA
Vall27-Asn195 (2281) CAGCGCATCGGCCGCCCTTGCTGGACATCCCCGCCGCA
Vall20-Ile201B (2239) CAGCGCATCGGCGGGGCCTTCCTGCACATCCCCGCCGCA
                    (2233) CAGCGCATCGGCCGCGCCTTCCTGCACATCCCCCGCGCA
 Val120-Ala204
                    (2239) CAGCCCATCGGCCGCCCTTCCTGCAGATCCCCCGCCGCA
 Val120-Ile201
 Vall20-Thr202 (2239) CAGCGCATCGGCCGCGCGCTTCCTGCACATCCCCCGCCGCA
Lys121-Val200 (2245) CAGCGCATCGGCCGCGCCTTCCTGCACATCCCCCGCCGCA
      Consensus (2281) CAGCGCATCGGCCGCGCCTTCCTGCACATCCCCCGCCGCA
                              2321
 Leu122-Ser199 (2291) TCCGCCAGGGCTTCGAGCGCCCTGCTGTAACTCGAGCG
                    (2321) TCCGCCAGGGCTTCGAGCGCGCGCGTCCTGTAACTCGAG--
(2279) TCCGGCAGGGCTTCGAGCGCGCGCGTGTAACTCGAGCG
 Val127-Asn195
Val120-Ile201B
 Vall20-Ala204 (2273) TCCSCCAGGCTTCGAGGGCTGCTGTAACTCGAG--Vall20-Ile201 (2279) TCCSCCAGGCTTCGAGCGGCTTGTAACTCGAG--
Vall20-Thr202 (2279) TCCGCCAGGGCTTCGAGCGCGCCCTGCTGTAACTCGAG--
Lys121-Val200 (2285) TCCGCCAGGGCTTCGAGCGCGCCCTGCTGTAACTCGAGCG
      Consensus
                    (2321) TCCGCCAGGGCTTCGAGCGCGCCCTGCTGTAACTCGAG
                              2361
 Leu122-Ser199 (2331) TGCT
 Val127-Asn195 (2359) ----
Val120-Ile201B
                    (2319) TGCT
 Val120-Ala204
                    (2311) ----
                   (2317) ----
 Val120-Ile201
 Val120-Thr202
                   (2317) ----
Lys121-Val200 (2325) TGCT
      Consensus
                   (2361)
```

WO 00/39303		1	15	1	65		PCT/US99/31272
Ile424-Ala433	(1)	_	CLOSES OF	enecesia.	reletava yes	unickticationacicinen	
Trp427-Gly431	(1)					1,70,210,210,010,010,010,01	
Gln422-Tyr435B						A Action article of Reli	
Arg426-Gly431	(1)					स्रोत्त्वत्रकः <u>)वीर्</u> देव्यक्ता	
Ile423-Met434	(1)					yalerizettirlelsivaten	
Gln422-Tyr435	(1)					natostánto ejercece	
Arg426-Lys432	(1)					अस्तार् <i>ोर्स्स</i> लंबान्स <mark>स्</mark>	
Arg426-Gly431B	(1)					n polyterine elegate by	
Asn425-Lys432	(1)					म् अस्त्रीक्रसंस्त्रहोस् <mark>र</mark> भूतिहा	
Consensus	(1)					AAGAGAGGGCTCT	
T1-424 N1-422		41			/FRED (SHIP)		80
Ile424-Ala433	(41)					११७४ हे हो भी झे हो हो दो दो दो <b>ले</b>	
Trp427-Gly431	(41)					ાનો સંવાદના માત્રોનો મુંદર્વે <b>ં વ્યક્કે</b> ટ	
Gln422-Tyr435B Arg426-Gly431	(41) (41)					કર મંતુમાર ભ્યક્ષોએ કેશોનો <b>લોક</b> મહોતાલ સભાવમાં ત્રામિક માને <b>ભીવો છે</b>	
Ile423-Met434	(41)						
Gln422-Tyr435	(41)					ા. ,લ,સલોઇસ્ટલને શ્રુપ જણાવેણી ૧૧ - ૧,લમાં લેકે છે કરો કે કે કે કોઇ	
Arg426-Lys432	(41)					กคายท่องจะอยู่จะเดียง	
Arg426-Gly431B	(41)					en in a clear and the ofele	111111111111111111111111111111111111111
Asn425-Lys432	(41)	•				: અનેલાં અનુવાલા જ હોલો <mark>છે</mark> છ	
Consensus	(41)	GTGTGCTGC				CTTCGTTTCGCC	
T1-424 D1-422	(01)	81					120
Ile424-Ala433	(81) (81)					१००मा । १४०० (तिस <b>स्</b> र	
Trp427-Gly431 Gln422-Tyr435B	(81)					त्रकारमञ्जूरशास्त्रकाराज्यान्तिः वर्षात्रकारमञ्जूरशास्त्रकाराज्यान्तिः	
Arg426-Gly431	(81)		***********			Solden a factor of the control of	
Ile423-Met434				***************************************		t machagis in the (stelle	
Gln422-Tyr435	(81)					has sometiment for	
Arg426-Lys432	(81)					्रकात्रीय अन्तर्भ स्थाप	
Arg426-Gly431B	(81)					system red a secrete	
Asn425-Lys432	(81)					कोर्यक्षाम् अद्गारम् सम्बद्धाः इतिहास	
Consensus	(81)					GTGTACTACGGC	
		121					160
Ile424-Ala433	(121)	State (1.15,048,640	वृत्तनाधान	7.5.000	કહ્યું વૃષ્ટ ફેર્ક જો	ઽઽઽૺ૽૱ૢ૽ઌ૽૽ૻઌ૿ૢઌ૽૽૽૱ઌૡ૽ૺૡ૿ૺ <b>ૢ</b>	
Trp427-Gly431	(121)	संस्थित प्रतिक तक्ष्रिक	Secole)	Assessed		WARRIED PROCE	
Gln422-Tyr435B	(121)	were continued	aporteres	Anafor	en singles are	का कार है। क्षेत्रकार माने की की क्ष	
Arg426-Gly431	(121)	P.Chrimage.isets	\$\$\te\e\n	ಕ್ಟಿಜರ್ಕ ಬ	अर्थ (श्रेंस्प्रोत्रे)	१७८८ (मिस्सेम्बर्ड स्ट्रि <b>ले</b> ड)	<b>22</b>
Ile423-Met434	(121)	व्यक्षाम पुरस्कितिहासि	fexetolog	अल्डिल्ड	क्षा के दिल्ला है।	रहार करने स्टब्स्ट्रेस्ट्रेस्ट्रेस्ट्रेस्ट्रेस्ट्रेस्ट्रेस्ट्र	200
Gln422-Tyr435	(121)	નીવાળ દાત્ર મેં ક્યાં છે છે	the contract	Migris	316 0 318 25	स्येर गुण्यस्य स्थान्त्रस्य	242A
Arg426-Lys432	(121)					ioti engearologa <b>eque</b>	
Arg426-Gly431B	(121)					equi, appropriatele	and the same of th
Asn425-Lys432	(121)		-,				
Consensus	(121)	CCCGTGTGG	AAGG	AGGCC	ACCACCA	CCCTGTTCTGCG	CCA 200
Ile424-Ala433	(161)		vojete(s)	्भभूस्टाट	atelaksissep.	१८१८ (कहाँ द्राजीहरू <mark>१८१८ (दर्श</mark> ी	
Trp427-Gly431	(161)	$\mathfrak{L}(\mathcal{E})^{\dagger}\mathfrak{L}(\mathfrak{L}(\mathcal{L})^{\dagger}\mathfrak{L}(\mathfrak{L})$	*:10,E(^.	r dutife	व्यक्ष (अस्त्री गुहरू)	randmateri/exts/rengij	TO S
Gln422-Tyr435B	(161)	are for the profession	<b>प्रश्नेत्र</b>	oprate (to	(415)4 <u>(616(61</u> )	(तर्मा, सन्दर्भ क्राप्तिक स्थान	5×86
Arg426-Gly431	(161)	elelejiteletelete	4:40de(c)	Carattella	elic hitelaler:	(स्ट्रां((e cs:\sv:y;76  <del>e</del> })(	≈ <b>∉</b> €
Ile423-Met434	(161)	stateste etalyas	except.	C 88417 (CS	१९४०-१५५०/१५ <u>१</u> ६ स	<i>च्याच्यां व्यक्तिको स्थाप</i>	511G
Gln422-Tyr435	(161)					refere e e e e e e e e e e e e e e e e e	
Arg426-Lys432	(161)	rickss/clolote/h	Material	wightele	feltigetiley:	(राष्ट्रीहराकाम्बरम <mark>्ह</mark> िस्	cv.(G
Arg426-Gly431B	(161)	بر <i>ي في ا</i> ما هاداسا	vividejeg	inglese:	Acheless	લ્લા) દેલગાદ <mark>ાસાયલન</mark> ો	el (G
Asn425-Lys432	(161)	वस्त्रम्यवाद्या	r:\e-(e)c/	profes	theres (class	ાલ્ફ્રેલ્સ્ક્સ્ટ્રેલ્સ્ફ્ <mark>રે</mark>	Selection of the select
Consensus	(161)		AGGC	CTACG	ACACCGA	GGTGCACAACGT	
Ile424-Ala433	(201)	201	le de or	∦e(\$!¢/i)	elejatasasia	्रेजुन्द्र स्ट्रीट के स्ट्रीट <b>्रा</b> स्ट्रीट हैं।	240 CAG
			B				

FIG. 4A

WO 00/39303		16 / 65	PCT/US99/31272
Trp427-Gly431	(201)	alete is contributed and contribute and contributed of the contributed	Mary Conference
Gln422-Tyr435B	(201		TOTAL CONTRACTOR
Arg426-Gly431	(201)		
Ile423-Met434	(201)		
Gln422-Tyr435	(201)		
Arg426-Lys432	(201)		
Arg426-Gly431B	(201)		
Asn425-Lys432	(201)		AND REAL PROPERTY.
Consensus	(201)		
Consensus	(201)	GGCCACCCACGCCTGCGTGCCCACCGACCCC	AACCCCCAG 280
Ile424-Ala433	(241)	uthousand cantle and each the trological the state of the	echiefer buch
Trp427-Gly431	(241)		
Gln422-Tyr435B	(241)	entreatring configuration for straightful and the configuration of the c	tiracrascias per etil
Arg426-Gly431	(241)		
Ile423-Met434	(241)		
Gln422-Tyr435	(241)		
Arg426-Lys432	(241)		
Arg426-Gly431B	(241)		
Asn425-Lys432	(241)		
Consensus	(241)	GAGATCGTGCTGGAGAACGTGACCGAGAACTT	CAACATGT
X1-404 31 400		281	320
Ile424-Ala433	(281)		
Trp427-Gly431	(281)		
Gln422-Tyr435B	(281)		
Arg426-Gly431	(281)		
Ile423-Met434	(281)		
Gln422-Tyr435	(281)		
Arg426-Lys432	(281)		
Arg426-Gly431B	(281)	etetingetingeningen inneten gebenketingheitetetinetetin	
Asn425-Lys432	(281)		
Consensus	(281)	GGAAGAACAACATGGTGGAGCAGATGCACGAG 321	GGACATCAT 360
Ile424-Ala433	(321)	wider, and a hactorophic chamber of payment can rede fa	
Trp427-Gly431	(321)		
Gln422-Tyr435B		क त्यार्थात संवत्त्र स्थाप देशाय । त्यार्थात विशेष विशेष विशेष विशेष स्थाप । विशेष संवत्त्र विशेष स्थाप ।	
Arg426-Gly431		Transcon in everythere ease of texture green by Province and the life	
Ile423-Met434	(321)		
Gln422-Tyr435	(321)		
Arg426-Lys432	(321)		
Arg426-Gly431B	(321)		
Asn425-Lys432	(321)		
Consensus	(321)		
71-404 33 400		361	400
Ile424-Ala433	(361)		
Trp427-Gly431	(361)	were mete content motion, ethicifel nell fre share little fill	
Gln422-Tyr435B	(361)	And the control of th	<del></del>
Arg426-Gly431	(361)	ार्था (१) । वाध्ययुक्तकार्वात्रे हे.देवलीस ८४०१३ विकित्तवीयोक्तरका अध्यक्ष स्वरूपका प्रकृ	
Ile423-Met434	(361)	हत है। या वार्डाह्मियमा १९ । स्वाया को में होने त्याद है पाप्रेस्ति है	
Gln422-Tyr435	(361)	WENT TO THE SERVICE OF THE PROPERTY OF THE PRO	
Arg426-Lys432	(361)	nervices elections are emphysical installes yeightestation	
Arg426-Gly431B	(361)	પ્રાજ્યાલુક મુજબાર હાથ છે. જે જે જે જે માન કર્યા છે.	
Asn425-Lys432	(361)	इति, तत्वा सार्गातम् दल्लाम् स्रोतिस्य स्थिते । स्थाना स्थाना स्थाना ।	
Consensus	(361)	ACCCCCTGTGCGTGACCCTGCACTGCACCAA	
T16424 B1-422		401	440
Ile424-Ala433	(401)	પ્રઅક્ષેત્રસંગ્રે(લે.પ્રાંતેમણા પ્રતિકારિલામાં પ્રાથમિક કર્યો હોય છે. કર્યો હોય કર્યો કર્યો હોય કર્યો કર્યા કર્યો કર્યા કર્યો કર્યા	
Trp427-Gly431	(401)	भूपति विकास मृत्युं व संस्थान स्मृत्युं है अने महास्थान स्मृत्युं के स्मृत्युं स्मृत्युं स्मृत्युं स्मृत्युं स	diversife (GE)
Gln422-Tyr435B	(401)	Helcie entimestationia feathlichtlandernerala inflestatet	क्रांत्रकार्यस्य होते

FIG. 4B

WO 00/39303		17 /	65	PCT/US99/31272
Arg426-Gly431	(401)	stretche in the forest profession and the services	redications, the separate and section of the first of the section	11 ( 2 ( 2 ) ) A
Ile423-Met434	(401)	Article Water Charles autominies	nero state in de minimización de con	
Gln422-Tyr435	(401)	בובלבי בין בין בין בין בין בין בין בין בין ב	ticicitions in the head of the interior	SVERTER .
Arg426-Lys432	(401)		grejestacie júziach neježekare tradest	
Arg426-Gly431B			underpreter antiquinterestrial elevations	0.0000
Asn425-Lys432	(401)		Activitate weaking the transfer action	
Consensus	(401)		AGCAGCAACTGGAAGGAGA	
		441		480
Ile424-Ala433	(441)	the planticial settle state of a state of a	हेर्द्रामध्यः ऋषावासीक्ष्यस्यहेर्द्राक्ष्यस्य	30.63
Trp427-Gly431	(441)	१ - २, २,धासद,संघोसकृतकृतकृत्रभाकार	Visitings at ethicing at 121 section 1917 c	Sitte 6 -
Gln422-Tyr435B	(441)	दः जीदावि विविद्याल्या विद्याले । अस्ति ।	प्रकासिक (असी ) सम्मूलिक प्रेजी को लाउन स	Migration .
Arg426-Gly431	(441)	equicing and enable administration for	तानीय द्वर १५५६१० मानीहरूत हुन १५५५५ ।	8 21 C
Ile423-Met434	(441)	ciclare, and any application servi-	A thresh textos tribre servensor in select	C-142-7g-
Gln422-Tyr435	(441)		germale professional and applications	
Arg426-Lys432			१८९५ <mark>। छोड् जिल्हा</mark> कारामधेर जिल्लाहरू स्टिस्	
Arg426-Gly431B	(441)	n, m, after, att, top v foat å kjost å t (a ) t k	१५१४मा वर्गमा हेर्न्स्य वर्गमा होत्रा १५५	COASS.
Asn425-Lys432	(441)		romici eticleranishiyikilennesher	
Consensus	(441)		CTGCAGCTTCAAGGTGAC	CACC
T1-404 N1-400	44011	481		520
Ile424-Ala433	(481)	resident in former taffer i transfer	letiterinifeletitehnitelefalalande	क्षेत्र कर्यके
Trp427-Gly431 Gln422-Tyr435B	(481)	Complete at a state statistical and	प्रकारित केल केले किरामान्त्रियां में मिला है।	West + 40
Arg426-Gly431	(481) (481)		esse distributed a telefolish in telefolish	
Ile423-Met434	(481)		terstellerstellstatelisteller ein sinkel	
Gln422-Tyr435	(481)		enterritaentelmit (en emilier leinemaelegitein geleit, more	
Arg426-Lys432	(481)		galgiest krye toatheite ta syc hes an gen terrane	
Arg426-Gly431B	(481)		felffejernelastychälfalelasson nei Militariose misternes als las son se	
Asn425-Lys432	(481)		ब्रियं के का दीवा प्रस्ता है हा हुए के के वार्ती	OCCUPATION OF THE PROPERTY OF
Consensus	(481)	AGCATCCGCAACAAGATC	CAGAAGGAGTACGCCCTG	<b>ΓΤ</b> (Τ
	, ,	521		560
Ile424-Ala433	(521)	ार्ड्य-क्रिकेट्राव्यक्ताम् <b>स्टब्स्ट्राव्यक्तर्यस्य स्टब्स्ट्र</b>	(कंटररहरूको त्यां सक्ता हु। क्षेत्र के स्व	
Trp427-Gly431	(521)		rodice of the production of the section	
Gln422-Tyr435B	(521)		Plant to Editorical Commission (Ar	
Arg426-Gly431	(521)		acceptiones data area Process	
Ile423-Met434	(521)		e egesti in itslige establessettistist et est	
Gln422-Tyr435	(521)	Solution of the state of the first	COLON CONT. WAS COMED TO SEE SEE	tristg
Arg426-Lys432	(521)	the hardware design deligners applied	(क. १५६६ कार कारत खेळकेला <mark>स</mark>	(a 5/4
Arg426-Gly431B	(521)		property a texast by a confidence of excess pass	
Asn425-Lys432	(521)	Acres in system of the classic rich	द्रावर्गमानुबन्धनान्। देशसम्बद्धानाः स्टब्स्	(१९१६)
Consensus	(521)		CCATCGACAACGACAACAC	CAG
T1-404 N1-400		561		600
Ile424-Ala433	(561)	ninerge state eten Konten nichte bie	sittita, meimitolokahitolohi aliteic	100 A C
Trp427-Gly431	(561)	a conte de la saciona de safet gent biblisse	kitikipa seleke protesiak etiko ya siriilio	9.649
Gln422-Tyr435B Arg426-Gly431	(561)		Outside rije in septiment lingstom Egiphelik	
Ile423-Met434	(561)		Systematic contests epitablishmanifeld	
Gln422-Tyr435	(561) (561)		क्षेत्रप्रेतक प्रदान के प्रतिकृतिक प्रदान के प्रतिकार के प्रतिकार के प्रतिकार के प्रतिकार के प्रतिकार के प्रतिक विकास के प्रतिकार के प्रति	
Arg426-Lys432	(561)		ejipageliaraciajimatel Agalasaciate ja 1990au 16 - Meli Primasa - 1963s	
Arg426-Gly431B	(561)		ujspaženja je užirči precih odana nasveje višša režanova oda v refi k eniška dišije vi	
Asn425-Lys432	(561)		s (skajdska (castajajajahhhilis) na silata Saakasta (. didiotash alis a saatata	
Consensus	(561)		CAACACCAGCGTGATCACC	
	(331)	601		640
Ile424-Ala433	(601)		rance that the entire that the contraction of the c	
Trp427-Gly431	(601)		tinte (chicicie le rouge) processione/e	
Gln422-Tyr435B			rakejeurejčjejergiacjejalicersićeje	
Arg426-Gly431			halaladelejejhaladeleti izatale	
Ile423-Met434			े देह , प्रियमिक्ट के का कार रहत है । इस देखीओ	

FIG. 4C

WO 00/39303		18 /	65	PCT/US99/31	272
Gln422-Tyr435	(601)	संस्टान्त्रसंस्ट्रीत्रहास्त्रीतात्त्र	letricit (698) lezensotal	PERMEMBER STREET, 19	
Arg426-Lys432	(601)				
Arg426-Gly431B	(601)				
Asn425-Lys432	(601)	सम्बद्धितार (स्थित स्थान	Chronitellic temperation	STABLICION TALCOTT A	
Consensus	(601)	GCCTGCCCCAAGGT	GAGCTTCGAGCC	CATCCCCATCCACT	
		641		680	
Ile424-Ala433	(641)	<i>पंदेश,</i> सन्दर्भक्षेत्रद्भक्षेत्रस्	स्थितः । । वहान्त्रस्य वहार्यस्य	engeftragefragefragetrage	
Trp427-Gly431	(641)	Maintais, el ciolo legal si al a	(દી <b>લ</b> ્લ મીજાલે (પ્લુપ્રફ્રેલ	enegatystejaketejasstrijeja	
Gln422-Tyr435B	(641)		(हरेट्राहरू स्वरूप का मेर्च हरे हैं। इस मेरिक ह	unkeptintepkeloutestallets	
Arg426-Gly431	(641)		,संबंध के श्रेष्ट्रस्थ <mark>(१५५४) है</mark>	cui, eztykfők nációlysárá jakk	
Ile423-Met434	(641)		Service and the service of the servi	<u>૽ૹઌૹૹૹ૽ૼૼૼૼૹૡઌઌઌૹ૽ૼૹ૽</u>	
Gln422-Tyr435	(641)				
Arg426-Lys432	(641)				
Arg426-Gly431B	(641)			<u>```````</u>	
Asn425-Lys432	(641)				
Consensus	(641)		GGCTTCGCCATC		
T1-404 N1-400	(01)	681		720	
Ile424-Ala433	(681)		GANCO NO SECONO	NAME OF THE PROPERTY OF THE PARTY OF THE PAR	
Trp427-Gly431 Gln422-Tyr435B		Secretarion de la company de l			
Arg426-Glv431	(681)	द्वाराज्यत् १ वर्ग्यक्त १४ म् स्टब्स्ट ४५५ ८ ह्व क्षर्यस्थिति वर्षायस्थ्यक्षर्यस्थिते वर्षे			
Ile423-Met434		grandly the first property of	e conte se que o como se se	The constitution of the	
Gln422-Tyr435	(681)				
Arg426-Lys432		\$257588585555555555555555555555555555555			
Arg426-Gly431B	(681)				
Asn425-Lys432	(681)		····		
Consensus	(681)				
		721		760	
Ile424-Ala433	(721)				
Trp427-Gly431	(721)		electric Control to the control of t	(ର୍ଜ୍ୟମଣ୍ଡାମ୍ବର, ପ୍ରଧିନ୍ତି	
Gln422-Tyr435B	(721)		સંદર્શન લેઇ (સોક્સેન્ટ) કરો છે છે.	्राह्मसङ्ग्रह्मसङ्ग्रह्मसङ्ग्रह्म	
Arg426-Gly431	(721)		****		
Ile423-Met434	(721)			<del></del>	
Gln422-Tyr435	(721)				
Arg426-Lys432	(721)				
Arg426-Gly431B	(721)				
Asn425-Lys432 Consensus	(721)	ne resettissigeistriffige	2 Marie 10 M	VE-02-22-2-2-2-2-2-2-2-2-2-2-2-2-2-2-2-2-	
Consensus	(721)	ACCGTGCAGTGCAC	CACGGCATCCGC	800	
Ile424-Ala433	(761)		dates made en kelt listede		
Trp427-Gly431	(761)				
Gln422-Tyr435B	(761)				
Arg426-Gly431	(761)				
Ile423-Met434	(761)	Most in an inference of the M			
Gln422-Tyr435	(761)	रत्या प्राप्तिक वृद्धां गणका समाने			
Arg426-Lys432	(761)	ระบาง การ์เปลี่ย เป็นผู้สารสาระกำระส			
Arg426-Gly431B	(761)	नक राष्ट्रको । सुक्ता अवस्था कार्यु	SEAT SEAL SHOOT IN SE	11.04654634636	
Asn425-Lys432	(761)	neterally partitional	Fredelator Meninakte	ार्टा दोर्डी द्वांताती रहे देश रहे भी	
Consensus	(761)		AACGGCAGCCTGG	CCGAGGAGGCGT	
		801	<u> </u>	840	
Ile424-Ala433	(801)				
Trp427-Gly431	(801)				
Gln422-Tyr435B	(801)				
Arg426-Gly431	(801)	सक्ष्मका के लेखाने में स्वति			
Ile423-Met434	(801)	etentistigniorefelelitzefelen			
Gln422-Tyr435	(801)	विक्रेडल (मानंदर्स्स) स्थातन			
Arg426-Lys432	(801)	ख्रुक्षम् दश्चमण्यत्।त्रान्युत्त <u>ात</u> ुत्त <u>ातुत्</u>	A STATE OF THE STATE OF CASE	अपन्य दाउँ (संस्कृत <b>ए</b>	

FIG. 4D

Arg426-Gly431B	(801)	
Asn425-Lys432	(801)	લાના ફુલ્માન હાલ્યાના પ્રાથમિક કાર્યા છે. જે કાર્યાના જે મહત્ર કોર્યાના કાર્યા કાર્યા કાર્યા છે. જે કાર્યા કોર્યા કોર્યા કાર્યા
Consensus	(801)	GGTGATCCGCAGCGAGACTTCACCGACAACGCCAAGACC
		841 880
Ile424-Ala433	(841)	entitier in tanger in a proposition of the second section of the sectio
Trp427-Gly431	(841)	
Gln422-Tyr435B	(841)	तमाद्राक्ष्यव्यक्षप्रभाइताद्रव्यक्षक्षक्ष्रकृतिकात्रक्षत्रक्षर्वाक्ष्यव्यक्षात्रक्ष्यक्ष्यव्यक्ष्यद्रवा
Arg426-Gly431	(841)	when the content to the fatter of the content of th
Ile423-Met434	(841)	
Gln422-Tyr435	(841)	क्षत्रक्षकार्यात्रवात्रवात्रवात्रवात्रवात्रवात्रवात्रव
Arg426-Lys432	(841)	क्षणीय प्रमान कन बन्दाका में तर भारत स्वतंत्र होता स्वतंत्र स्वतंत्र में स्वतंत्र स्वतंत्र स्वतंत्र स्वतंत्र स
Arg426-Gly431B	(841)	teres the new as alternacie, retranscious estancians carreternos, estante mension
Asn425-Lys432	(841)	באיצי בה בבול ביו בלכול בקיפה (ביובה לבול בלו ליורן ס בא לובינה ולה לבול הליב באלכול בל
Consensus	(841)	ATCATCGTGCAGCTGAAGGAGAGCGTGGAGATCAACTGCA
	(0.12)	881 920
Ile424-Ala433	(881)	ान्द्रवादिहर्ष ते व्यक्ति व्यवस्थानम् । स्वत्यक्ति व्यक्तिम् स्वत्यक्ति । स्वत्यक्ति । स्वत्यक्ति । स्वत्यक्ति
Trp427-Gly431	(881)	nen selente la tritala la appenta calentralem esta fentalmit la la
Gln422-Tyr435B	(881)	
Arg426-Gly431	(881)	รได้เรียนให้เล่นสายเหมือน เขาที่เหมือนเพาะเล่น เล่าสามารถสายเรียนให้เล่นสายเล่นสายเหมือน ผู้เรียน
Ile423-Met434		e (Carteen (Cartenesias), inventingal infahritah Heli inne Yesesteber 16.
Gln422-Tyr435	(881)	whose the transfer of the foregreen is the figure of the state of the
	(881)	age to test, respect to the strategies of the strategies of the state of the strategies and the strategies of the strate
Arg426-Lys432	(881)	क्षेत्रीय, स्था, १८ प्राप्त-प्रमुख्या के एक प्रथम, अस्ति । ता प्रथम क्षेत्रक विकास स्थाप विद्या विद्या विद्या
Arg426-Gly431B	(881)	cjajo teknis i josekus, eneroj prefesuojekus i jednikeniniste interioriteks, eta i judicioriteks, ede
Asn425-Lys432	(881)	अत्राहरक्षेत्र्रात्र्वे स्थापन्त्रात्स्य राज्ये प्रेरेष्ट वर्षः वर्षः स्थापन्तर्वात्रः स्थापन्तर्वात्रः स्थापन
Consensus	(881)	CCCGCCCAACAACACCCCGCAAGAGCATCACCATCGG
73 404 73 477		921 960
Ile424-Ala433	(921)	elacticiació o delación para felar es socialmente de destinar alicitific
Trp427-Gly431	(921)	क्षर्रहास्त्रहारा । त्रहितास्त्रिक्ताराज्यकारणदार्थाकार्यक त्राताराज्यसम्बद्धाः व्यवस्थितसम्
Gln422-Tyr435B	(921)	नाम्यकारम् स्थात्राम्यवराद्यः । सम्यवनाम् अनुसर्वात्रम् । तामान्यस्य क्षात्रम् । स्थान्यस्य ।
Arg426-Gly431	(921)	後に対けはははなる日本はは、それにこれなるかにいかない。 きはとはかけらいのことにはななる
Ile423-Met434	(921)	अवदेशां स्टब्स्य होते । इंग्लेक्स्य वर्षे अधिक विदेशते स्टब्स्य होते होते होते होते होते होते होते होते
<b>Gln422-Tyr43</b> 5	(921)	wish the inferior colors of the state of the
Arg426-Lys432	(921,)	interview (and the first of the independent of Course and content of the dynamic of the content
Arg426-Gly431B	(921)	कर के में कार को देखिए हैं है। अने कि कार के कि कार के कि कार की कि
Asn425-Lys432	(921)	લીકા લીકા કરવાના જોવા જોવાર જેમાં મામે કાંમાને કે જેવા કરતા હોય કાર્યકાર <mark>કરને પ્રદેશની પાસે કરે</mark> વાર્યકાર્યકાર છ
Consensus	(921)	CCCCGGCCGCCTTCTACGCCACCGGCGACATCATCGGC
		961 1000
Ile424-Ala433	(961)	्रेष्ट्रय १ रे देश हे स्थाप्त हे से स्थापन हो स्थापन स्थापन स्थापन स्थापन स्थापन स्थापन स्थापन स्थापन स्थापन स
Trp427-Gly431	(961)	cause the activities of the firm of grane integrations the violate (or equipment of extension)
Gln422-Tyr435B	(961)	entsymme of each mount on her mentally have the printing entering the transfer
Arg426-Gly431	(961)	And were a respectively wealth in opposite from a desperior material. In
Ile423-Met434	(961)	Alternative to the properties of the properties
Gln422-Tyr435	(961)	Spirit Dagger of the good Control of Spirits and a few sections of the control of
Arg426-Lys432	(961)	Record (1946) of the production of proposal articulation record and local desired physical production of the production
Arg426-Gly431B	(961)	कर्मकरम्मता अस्ति पुर प्रदर्भः । त अनेरचाति । विक्रियं कर्मकर्मम् । अस्ति । अस्ति । विक्रियं । विक्रियं । विक्रिय
Asn425-Lys432	(961)	mingkegelinger in eine mingeten in die eine kannen sich eine in ereiche er er eine keine bei ein die der eine die eine eine eine eine eine eine
Consensus		CACATCCCCCA CCCCCA CTCCA A CATTCA CCCCCCA CACA
consensus	(961)	GACATCCGCCAGGCCCACTGCAACATCAGCGGCGAGAAGT 1001 1040
Ile424-Ala433	(1001)	
	(1001)	electrite appelitelete a etapotetetrostassión editeletratete, gét Minicie
Trp427-Gly431 Gln422-Tyr435B	(1001)	यान्त्रभागतः । स्वतः वर्षातः । वर्षात्रभागति । स्वतः । प्रतः । प्रतः । स्वतः । स्वतः । स्वतः । स्वतः । स्वतः व
	(1001)	ejejitžja dytikitijt čatenostajtestiičitaju, elepnojujeje bitateizin eleptaiole
Arg426-Gly431	(1001)	elminate reit erriet tammetrateckatethemeter meter mister etritation infortiere e
Ile423-Met434	(1001)	egentyste isakys sega, g. m. nesthologees kohym, ey dan, media i drejega kinginyas en se
Gln422-Tyr435	(1001)	त्रीक्षक्रमेतः इत्रीक्षमितं कृष्णक्षीक्षेत्रमैत्वेद्वक्षमितिकुष्टिकार एम्पर्किकार्यद्वकृष्टिक्षायातिकृष्ट
Arg426-Lys432	(1001)	ब्युट्ट्रेस्स्स्रेट्स्स्स्स्रेट्र्स्स्स्रेट्र्स्ट्र्स्स्स्स्रेस्स्स्स्स्स्स्रेस्स्स्रेस्स्स्रेस्स्स्रेस्स्स्रेस्स्स्
Arg426-Gly431B	(1001)	equalption they be being an antither by the being contracted by the filter being the being a second
Asn425-Lys432	(1001)	ethe yakaye, a aher siye heye ilder a serve siyê bil. Er biye yê biyeye ye stalak hike ta siyasire. Ç

FIG. 4E

PCT/US99/31272

Consensus (1001) GGAACAACACCCTGAAGCAGATCGTGACCAAGCTGCAGGC 1041 Ile424-Ala433 (1041)อยุคมายสุดผลเลยเลียงสมอยที่เรียงให้เรื่องสีสัตร์แล่งสุดใหล่เห็นได้สายเลยในการ Trp427-Gly431 (1041) edestretus acepetete susprient/estratations to sensi a discussioni discontrate testrite (1041) Nether kern engagelete digitabilitation betrestive etablete ingeneeste Gln422-Tvr435B Arg426-Gly431 (1041) months in the medical supplies and the step in the language successive Ile423-Met434 (1041) To stop the letter state of the section of the state of the sta Gln422-Tyr435 (1041)Arg426-Lys432 (1041) tight about a thoromorphy park complicates general particular in Elica of America (1970) for the Professional (1041) चित्रकाताको विकासकारिकामुख्यानुस्थानुस्थानुस्थानुस्थानुस्थानुस्थानुस्थानुस्थानुस्थानुस्थानुस्थानुस्थानु Arg426-Gly431B Asn425-Lys432 (1041) Machen a ciclopart appropriate property of the party of the city of the Consensus 1081 Ile424-Ala433 (1081) त्यानुः (त्याद्व (देशको के देशको क्षुत्रकारोको देशको अध्योत्व स्थापन्त । व्यक्तिको विकास करें विकास करें के कि Trp427-Gly431 contra termiconomia el mententenno el termicontenno de la contra (1081) Gln422-Tyr435B (1081) efetsteingeleten jersjelengtengsloppidenten elementeleficht internetive e e Arg426-Gly431 (1081) Get-infrible met telagotet dieget alla bet de transferior de la conference de la con Ile423-Met434 (1081) meinteletete in volgtetile in het te keine volgte volgtetie in volgtetie in Gln422-Tyr435 (1081)वार्याका होताम प्रमानेक में विविधिक में प्रमानिक व्यक्ति होता है व्यक्ति होता है है है । यह से प्रमान के महिद्द ejantanarian ijak ripnjergesterijang sammanaritebele and Arg426-Lys432 (1081) Arg426-Gly431B (1081) CONSTRUCTION OF STATE OF TO SELECT A STATE ASSESSED THE ASSESSATION OF THE STATE OF THE ASSESSATION OF THE ASSESSATION (1081) The state court the completeness integral and another training the defect Asn425-Lys432 Consensus (1081)GGCGGCGACCCCGAGATCGTGATGCACAGCTTCAACTGCG 1160 स्टरक्टर्स्टर है अन्तरभाषा मान्यून प्रमुक्त स्वार्ड करें के स्वार्ड करें के स्वर्ड के स्वर्ड के स्वर्ड के स्वर Ile424-Ala433 (1121)Trp427-Gly431 (1121) व्यववर्गन्त्रकार्वकार्वकार्यक्षिक्ष्यां व्यवस्थित्वा Gln422-Tyr435B alegeren felaban ihr inde fallentille fartentille artenpeter fir belogetor gentellen i (1121)Arg426-Gly431 (1121)અલ્લાહાલ્યા, કોર્યક્ષ, દુવાનો, સાંદુર્ન કેડ્ડાલાકો અલ્લોને ઉત્તર દેવના હોય જેવા કાર્યક્રોનો દુવા જાઈ Ile423-Met434 (1121)(1121) desire bases es referências de designation esta que accellança en Gln422-Tyr435 Arg426-Lys432 (1121) resistantetra instribute paga eleterativa benegativa de la Applicação Para el Constituição de la Cons Arg426-Gly431B (1121) the the end of the end of the property of the energy of the end of the Asn425-Lys432 (1121) Andrew State Committee of the programme forms to engine programming Consensus (1121) GCGGCGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAA 1161 1200 Ile424-Ala433 egate majanek an elitakigayankegania kakatani minintelitik olitaki pipanake<u>zaani kaki</u> ma<u>ta</u> (1161) Trp427-Glv431 (1161) Gln422~Tyr435B (1161)the angle of the companies of the series of the property of th Arg426-Gly431 (1161)्रमा ६ १ वर १८० वर्षा क्रिकेट प्राप्त हुन्मक होता हर हुन है स्वरूप राज्य हुन्<mark>न स्पार्त है स्वरूप स्वरूप स्वरूप</mark> Ile423-Met434 (1161) ा अञ्चलक्षेत्रक व्यामक्षेत्रक अञ्चलक्ष्यक स्थाप । १ व व्यवस्थित व्यक्तिक वि Gln422-Tyr435 (1161) where the continuous residual for the magniference that Arg426-Lys432 (1161) with the antique procedure of the Arg426-Gly431B (1161) march loss outsidants from the control of the Action to be the feeting Asn425-Lys432 (1161) Proceeding Companies and the procedure of the proc Consensus (1161) CAGCACCTGGAACAACACCATCGGCCCCAACAACACCAAC Ile424-Ala433 (1201) Ideas the new drophing to properly properly of the control eten kanalana irki projetoja mekinemiyletolmie ani 10,0 men rakvanja more j Trp427-Gly431 (1201) Gln422-Tyr435B क्षेत्रः क्षेत्रको हेमा क्षेत्रकिते कार्यसार (देश हेस्सी हेस्सान के सकतन्त्रको कार्य------(1201) Arg426-Gly431 (1201) विविधान विकास विकास विकास विकास के वित्र के विकास Ile423-Met434 (1201) June 19 Germeters Conference of particle of protection of the content of t Gln422-Tvr435 (1201) step 100 few 100000 revenue of the section o Arg426-Lys432 (1201) upor fate lapando fat tra e triplo accorden Jacob para para proportion (4 Arg426-Gly431B (1201) epitrati influenci protesionni fallomi deput desimble de sindici de la companya de la com (1201) etale escalator estaleja escalator estaleja etaleta estaleja estaleta estaleta estaleta estaleta esperi Asn425-Lys432

Consensus

FIG. 4F

(1201) GGCACCATCACCCTGCCCTGCCGCATCAAGCAGATCATCA

21 / 65

Ile424-Ala433	(1240)	द्रविल्स्तर्द् ह - स्वर्थनगर स्वर्थर द्रावा हरित हरित हरित हरित हरित हरित हरित हरित
Trp427-Gly431	(1241)	ABCGCTGG द्वांबुद्धवाल्यक्ष्मान्यकार्यक्षात्रकार्यकार्यकार्यकार्यकार्यकार्यकार्यकार
Gln422-Tyr435B	(1234)	ECECEC
Arg426-Gly431	(1241)	ःश्रि <b>CGCGGC</b> वृद्धानिविक्षेत्रकातांविक्षकेत्रकेत्रवात्ताताः । १९८१ वर्षः १००० वर्षः । १०००
Ile423-Met434	(1237)	etelejejete Wale wate jeje je je jeje it je jestilik e je
Gln422-Tyr435	(1234)	
Arg426-Lys432	(1241)	A CGCGGC HAGAA DINNELS COLUMN
Arg426-Gly431B	(1241)	TO CGCGGCA TO HELEN WAS CONTROL TO THE TOTAL OF THE TOTAL
Asn425-Lys432	(1241)	are GCSCC hithardeleietan ejagin (Legate, par jedetige ente
Consensus	(1241)	AC GGCGGCAAGGCCATGTACGCCCCCCCCATCCG
		1281 1320
Ile424-Ala433	(1269)	दाकाद्रकाद्रमास्त्रकामभग्नीन <b>्</b> रिकासकाम् मञ्जलकारीत्रकारीयानाम् संभावनः दावान् अत्यक्षमः । रीव
Trp427-Gly431	(1281)	वृद्धिक (स्थापक्षका द्विष्ठ प्रकृष्टिक में स्थापक क्षेत्र । वर्षे प्रकृष्टिक व्यक्ति । वर्षे
Gln422-Tyr435B	(1257)	स्तित्वा करेत्राच्या मृद्धा <b>स्ति स्ति स्ति</b> स्टिस स्ति स्ति स्ति स्ति स्ति स्ति स्ति स्
Arg426-Gly431	(1281)	राहारा र राजेद्देशमध्यान् द्रवानां वास्ताने वास्ताने वास्तान वास्तान वास्तान वास्तान
Ile423-Met434	(1263)	क्षा प्रकार का प्रमाण का
Gln422-Tyr435	(1257)	
Arg426-Lys432		ingletar gladijski officielovik (in likularian) i pravolar majarovi jelovija mjerni se
Arg426-Gly431B	(1281)	missional elistric succession de la company de la comp
Asn425-Lys432	(1281)	લંશાલા પ્રાપ્તિ કાર્યાં દેવિની પશ્ચિક છે. તે કર્યું હોંકાલ કરે કર્યા છે. તે કર્યા હોંકાલ કરે કર્યા હોંકાલ કરે
Consensus	(1275)	And the test of the confidence of the second section of the confidence of the confid
Consensus	(1281)	CGGCCAGATCCGCTGCAGCAGCAACATCACCGGCCTGCTG 1321 1360
Ile424-Ala433	(1309)	1000
Trp427-Gly431	(1321)	૧૩ માલા માર્ગ્યું કે પાલા આપી લોકો વાર્ષ કે માર્ચ્યું કે કે કે બે આવા તાલા આ મોટા કે પાલા કે અને કરો છે. આ દે ભાગમાં કે આ આપી કે માર્ચ્યું કે માર્ચ્યું કે કે કે કે માર્ચિક માર્ચિક માર્ચિક માર્ચિક માર્ચિક માર્ચિક માર્ચિક મ
Gln422-Tyr435B	(1297)	
Arg426-Gly431	(1321)	en in the first of the management of the first of the fir
Ile423-Met434	(1303)	च्यात्रक्षात्रकारम् । स्वत्रकारम् । स्वत्रकारम् । स्वत्रकारम् । स्वत्रकारम् । स्वत्रकारम् । स्वत्रकारम् । स्वत स्वत्रकारम् । स्वत्रकारम्
Gln422-Tyr435	(1297)	
Arg426-Lys432	(1321)	લાકાર્ય કે તાલ કર્યું છે. જિલ્લાએ જે ઉત્તર એ જાણ કે કે કે કે કે કારણ કરવા છે. તેના જ પ્રયોગ હતા, તેને કે કે કે જ્યારા તે તેના તેના તેના કે
Arg426-Gly431B	(1321)	o upanteletajo ist <b>erija je i</b> ki tejo iš podije intervensko najvinje si vina vise vise i je intervenski koji intervise i je in
Asn425-Lys432	(1315)	Christian in the control of the cont
20 275452	(1313)	
Consensus	(1321)	
Consensus	(1321)	CTGACCCGCGACGGCGGCAAGGAGATCAGCAACACCACCG
		CTGACCCGCGACGGCGAAGGAGATCAGCAACACCACCG 1361 1400
Ile424-Ala433	(1349)	CTGACCCGCGACGGCGGCAAGGAGATCAGCAACACCACCG 1361 1400
Ile424-Ala433 Trp427-Gly431	(1349) (1361)	CTGACCCGCGACGGCGGCAAGGAGATCAGCAACACCACCG 1361 1400
Ile424-Ala433 Trp427-Gly431 Gln422-Tyr435B	(1349) (1361) (1337)	CTGACCCGCGACGGCGGCAAGGAGATCAGCAACACCACCG 1361 1400  Additional of the Control of
Ile424-Ala433 Trp427-Gly431 Gln422-Tyr435B Arg426-Gly431	(1349) (1361) (1337) (1361)	CTGACCCGCGACGGCGGCAAGGAGATCAGCAACACCACCG 1361 1400  Marine of the control of the
Ile424-Ala433 Trp427-Gly431 Gln422-Tyr435B Arg426-Gly431 Ile423-Met434	(1349) (1361) (1337) (1361) (1343)	CTGACCCGCGACGGCGGCAAGGAGATCAGCAACACCACCG 1361 1400  Additional of the control of
Ile424-Ala433 Trp427-Gly431 Gln422-Tyr435B Arg426-Gly431 Ile423-Met434 Gln422-Tyr435	(1349) (1361) (1337) (1361) (1343) (1337)	CTGACCCGCGACGGCGGCAAGGAGATCAGCAACACCACCG 1361 1400  Addition of the control of th
Ile424-Ala433 Trp427-Gly431 Gln422-Tyr435B Arg426-Gly431 Ile423-Met434 Gln422-Tyr435 Arg426-Lys432	(1349) (1361) (1337) (1361) (1343) (1337) (1361)	CTGACCCGCGACGGCGGCAAGGAGATCAGCAACACCACCG 1361 1400  Addition of the control of th
Ile424-Ala433 Trp427-Gly431 Gln422-Tyr435B Arg426-Gly431 Ile423-Met434 Gln422-Tyr435 Arg426-Lys432 Arg426-Gly431B	(1349) (1361) (1337) (1361) (1343) (1337) (1361) (1361)	CTGACCCGCGACGGCGGCAAGGAGATCAGCAACACCACCG 1361  1400  Astronomy Control of Con
Ile424-Ala433 Trp427-Gly431 Gln422-Tyr435B Arg426-Gly431 Ile423-Met434 Gln422-Tyr435 Arg426-Lys432 Arg426-Gly431B Asn425-Lys432	(1349) (1361) (1337) (1361) (1343) (1337) (1361) (1361) (1355)	CTGACCCGCGACGGCGGCAAGGAGATCAGCAACACCACCG 1361  1400  Astrice of Control of Co
Ile424-Ala433 Trp427-Gly431 Gln422-Tyr435B Arg426-Gly431 Ile423-Met434 Gln422-Tyr435 Arg426-Lys432 Arg426-Gly431B	(1349) (1361) (1337) (1361) (1343) (1337) (1361) (1361)	CTGACCCGCGACGGCGGCAAGGAGATCAGCAACACCACCG 1361 1400 1400 1400 1400 1400 1400 1400 14
Ile424-Ala433 Trp427-Gly431 Gln422-Tyr435B Arg426-Gly431 Ile423-Met434 Gln422-Tyr435 Arg426-Lys432 Arg426-Gly431B Asn425-Lys432 Consensus	(1349) (1361) (1337) (1361) (1343) (1337) (1361) (1361) (1355) (1361)	CTGACCCGCGACGGCGGCAAGGAGATCAGCAACACCACCG 1361 1400 1400 1400 1400 1400 1400 1400 14
Ile424-Ala433 Trp427-Gly431 Gln422-Tyr435B Arg426-Gly431 Ile423-Met434 Gln422-Tyr435 Arg426-Lys432 Arg426-Gly431B Asn425-Lys432 Consensus	(1349) (1361) (1337) (1361) (1343) (1337) (1361) (1361) (1355) (1361) (1389)	CTGACCCGCGACGGCGGCAAGGAGATCAGCAACACCACCG 1361 1400 1510 1510 1510 1510 1510 1510 151
Ile424-Ala433 Trp427-Gly431 Gln422-Tyr435B Arg426-Gly431 Ile423-Met434 Gln422-Tyr435 Arg426-Lys432 Arg426-Gly431B Asn425-Lys432 Consensus Ile424-Ala433 Trp427-Gly431	(1349) (1361) (1337) (1361) (1343) (1337) (1361) (1361) (1355) (1361) (1389) (1401)	CTGACCCGCGACGGCGGCAAGGAGATCAGCAACACCACCG 1361 1400 1516 1516 1516 1516 1516 1516 1516 15
Ile424-Ala433 Trp427-Gly431 Gln422-Tyr435B Arg426-Gly431 Ile423-Met434 Gln422-Tyr435 Arg426-Lys432 Arg426-Gly431B Asn425-Lys432 Consensus Ile424-Ala433 Trp427-Gly431 Gln422-Tyr435B	(1349) (1361) (1337) (1361) (1343) (1337) (1361) (1361) (1355) (1361) (1389) (1401) (1377)	CTGACCCGCGACGGCGGCAAGGAGATCAGCAACACCACCG 1361  1400  1400  1400  1400  1400  1400  1400  1400  1400  1400  1400  1400  1400  1400  1400  1400  1400  1400  1440
Ile424-Ala433 Trp427-Gly431 Gln422-Tyr435B Arg426-Gly431 Ile423-Met434 Gln422-Tyr435 Arg426-Lys432 Arg426-Gly431B Asn425-Lys432 Consensus Ile424-Ala433 Trp427-Gly431 Gln422-Tyr435B Arg426-Gly431	(1349) (1361) (1337) (1361) (1343) (1337) (1361) (1361) (1355) (1361) (1389) (1401) (1377) (1401)	CTGACCCGCGACGGCGGCAAGGAGATCAGCAACACCACCG 1361  1400  1440  1440
Ile424-Ala433 Trp427-Gly431 Gln422-Tyr435B Arg426-Gly431 Ile423-Met434 Gln422-Tyr435 Arg426-Lys432 Arg426-Gly431B Asn425-Lys432 Consensus Ile424-Ala433 Trp427-Gly431 Gln422-Tyr435B Arg426-Gly431 Ile423-Met434	(1349) (1361) (1337) (1361) (1343) (1337) (1361) (1355) (1361) (1389) (1401) (1377) (1401) (1383)	CTGACCCGCGACGGCGGCAAGGAGATCAGCAACACCACCG 1361  1400  1440  1440  1440  1440  1440  1440
Ile424-Ala433 Trp427-Gly431 Gln422-Tyr435B Arg426-Gly431 Ile423-Met434 Gln422-Tyr435 Arg426-Lys432 Arg426-Gly431B Asn425-Lys432 Consensus Ile424-Ala433 Trp427-Gly431 Gln422-Tyr435B Arg426-Gly431 Ile423-Met434 Gln422-Tyr435	(1349) (1361) (1337) (1361) (1343) (1337) (1361) (1361) (1355) (1361) (1389) (1401) (1377) (1401) (1383) (1377)	CTGACCCGCGACGGCGGCAAGGAGATCAGCAACACCACCG 1361  1400  1
Ile424-Ala433 Trp427-Gly431 Gln422-Tyr435B Arg426-Gly431 Ile423-Met434 Gln422-Tyr435 Arg426-Lys432 Arg426-Gly431B Asn425-Lys432 Consensus Ile424-Ala433 Trp427-Gly431 Gln422-Tyr435B Arg426-Gly431 Ile423-Met434 Gln422-Tyr435 Arg426-Lys432	(1349) (1361) (1337) (1361) (1343) (1337) (1361) (1361) (1355) (1361) (1377) (1401) (1383) (1377) (1401)	CTGACCCGCGACGGCGGCAAGGAGATCAGCAACACCACCG 1361  1400  1
Ile424-Ala433 Trp427-Gly431 Gln422-Tyr435B Arg426-Gly431 Ile423-Met434 Gln422-Tyr435 Arg426-Lys432 Arg426-Gly431B Asn425-Lys432 Consensus Ile424-Ala433 Trp427-Gly431 Gln422-Tyr435B Arg426-Gly431 Ile423-Met434 Gln422-Tyr435 Arg426-Lys432 Arg426-Gly431B	(1349) (1361) (1337) (1361) (1343) (1337) (1361) (1361) (1355) (1361) (1389) (1401) (1377) (1401) (1383) (1377) (1401) (1401)	CTGACCCGCGACGGCGGCAAGGAGATCAGCAACACCACCG 1361  1400  1440  1
Ile424-Ala433 Trp427-Gly431 Gln422-Tyr435B Arg426-Gly431 Ile423-Met434 Gln422-Tyr435 Arg426-Lys432 Arg426-Gly431B Asn425-Lys432 Consensus Ile424-Ala433 Trp427-Gly431 Gln422-Tyr435B Arg426-Gly431 Ile423-Met434 Gln422-Tyr435 Arg426-Lys432 Arg426-Gly431B Asn425-Lys432	(1349) (1361) (1337) (1361) (1343) (1337) (1361) (1361) (1355) (1361) (1401) (1383) (1401) (1401) (1401) (1401) (1401) (1401) (1395)	CTGACCCGCGACGGCGGCAAGGAGATCAGCAACACCACCG 1361  1400  1
Ile424-Ala433 Trp427-Gly431 Gln422-Tyr435B Arg426-Gly431 Ile423-Met434 Gln422-Tyr435 Arg426-Lys432 Arg426-Gly431B Asn425-Lys432 Consensus Ile424-Ala433 Trp427-Gly431 Gln422-Tyr435B Arg426-Gly431 Ile423-Met434 Gln422-Tyr435 Arg426-Lys432 Arg426-Gly431B	(1349) (1361) (1337) (1361) (1343) (1337) (1361) (1361) (1355) (1361) (1389) (1401) (1377) (1401) (1383) (1377) (1401) (1401)	CTGACCCGCGACGGCGGCAAGGAGATCAGCAACACCACCG 1361  1400  1
Ile424-Ala433 Trp427-Gly431 Gln422-Tyr435B Arg426-Gly431 Ile423-Met434 Gln422-Tyr435 Arg426-Lys432 Arg426-Gly431B Asn425-Lys432 Consensus Ile424-Ala433 Trp427-Gly431 Gln422-Tyr435B Arg426-Gly431 Ile423-Met434 Gln422-Tyr435 Arg426-Lys432 Arg426-Gly431B Asn425-Lys432 Consensus	(1349) (1361) (1337) (1361) (1343) (1337) (1361) (1361) (1355) (1361) (1389) (1401) (1377) (1401) (1383) (1377) (1401) (1401) (1395) (1401)	CTGACCCGCGACGGCGGCAAGGAGATCAGCAACACCACCG 1361  1400  1511  1
Ile424-Ala433 Trp427-Gly431 Gln422-Tyr435B Arg426-Gly431 Ile423-Met434 Gln422-Tyr435 Arg426-Lys432 Arg426-Gly431B Asn425-Lys432 Consensus Ile424-Ala433 Trp427-Gly431 Gln422-Tyr435B Arg426-Gly431 Ile423-Met434 Gln422-Tyr435 Arg426-Lys432 Arg426-Gly431B Asn425-Lys432	(1349) (1361) (1337) (1361) (1343) (1337) (1361) (1361) (1361) (1361) (1377) (1401) (1383) (1377) (1401) (1401) (1401) (1401) (1401) (1401) (1401) (1401)	CTGACCCGCGACGGCGGCAAGGAGATCAGCAACACCACCG 1361  1400  1

FIG. 4G

C1 - 400 M 405 N		
Gln422-Tyr435B	(1417)	equipment material equipment all productions and the production of
Arg426-Gly431	(1441)	. जोट तम्मान्त्रम् (हार्यमिनेन्द्रोनीत हार १०१६८) । अस्मान्त्रात्रात्रम् अस्तान्त्रात्रात्रात्रात्रात्रात्रात्
Ile423-Met434	(1423)	ांगोल्या प्रातिस्तित्वालक्ष्मीत्वे । भवेत् वर्षा प्रातिस्य प्रिया नामवन्त्रत्ये। भवेत् स्त्राव्यस्य प्रात्स्य
Gln422-Tyr435	(1417)	cloud, with the paper and the property of the
Arg426-Lys432	(1441)	esa ta la impalable (emblica serale) e (emblic este fect e la pablica papillo serales antice
Arg426-Gly431B	(1441)	n extra regalatin e ancietu (Ljate ja e in e in rygnasekete fetifatojata ja kiki jaku jaku jaku jaku jaku jaku
Asn425-Lys432	(1435)	द्याया स्वाह्मा क्षेत्री का संस्थानीत हिल्ला है। इस्तान इस्ति का स्वाह्मा हो है। इस्ति है। इस्ति है।
Consensus	(1441)	CCCCTGGGCGTGGCCCCACCAAGGCCAAGCGCCGCGTGG
		1481 1520
Ile424-Ala433	(1469)	भारतिकारमान्यकार 'कार्यक्रिकेशकादिक्षिक्ष दिल्ला देशकार्थिकार क्षित्र दिल्लाकार होता है।
Trp427-Gly431	(1481)	men innepeter jennen ysperentetet et einten hen hit innepetet et eletter har hit
Gln422-Tyr435B	(1457)	namendagestatestatestatestatestates and participates and antital and antital and antital and antital and antital antital antital and antital a
Arg426-Gly431	(1481)	भागमान्या जन्मान्यकारेशा स्टिब्ब्यियानम् । स्ट्रान्यकारस्य विश्वार्यस्य
Ile423-Met434	(1463)	तातर अर्था होत्यो। काच्या प्रकृतिक सम्भवित अर्थान कर्या आया प्रतिविद्यास्य क्रिकेट अर्थ
Gln422-Tyr435	(1457)	भारताह , भारताह , हा द्वारा कर कर होते कर होते कर होता है। एक कर मानाह हात होता है। हात होता है कि के के के के
Arg426-Lys432	(1481)	antegraph in teachastate, continue class contributions feld elegensisten
Arg426-Gly431B	(1481)	तवान्त्रपद्भवात्तात्वात्रकात्रात्रवातित्रवातः दश्यात्वा (व्यव्यवस्थान्त्रवाद्भवस्त्रवितः
Asn425-Lys432	(1475)	ताकारेक्षात्र श्रेष्टावाकारेक्षात्र । देखावाक्षिकार । स्वयंत्र श्रेष्टा । स्वयंत्र । स्वयंत्र । स्वयंत्र । स्व
Consensus	(1481)	TGCAGCGCGAGAAGCGCGCCGTGACCCTGGGCGCCATGTT
		1521 1560
Ile424-Ala433	(1509)	्या । व्यवस्थाः वा वा वा वा व्यवस्थाने ।
Trp427-Gly431	(1521)	५५ मार्ग्याहरेल । जनके प्रतासकात सिक्षेत्र मा रूपना शिक्ष पृक्षाहरेगी मेनोक्स वर्षी प्रकाश स्थित
Gln422-Tyr435B	(1497)	અભ્યાન કર્યાં છે. આ કાર્યા ભારત કરો કરો છે. જે જો મોળ મોળ મોળ કરી જો છે. જો જ
Arg426-Gly431	(1521)	र देशक्षक व्यवकारम् रहे। १७४४ व्यवस्थात्त्रः १०६४ (१.५०) वर्षः १९६८) वर्षः १०१४ वर्षः १०१४ वर्षः १०१४ वर्षः
Ile423-Met434	(1503)	. में १४८ व दर्शने पर प्रकास व्यवस्थित विद्यान है। विद्यान विद्यान के स्थान मुख्ये के विद्यान कि विद्यान के वि
Gln422-Tyr435	(1497)	the engine is the continue of the following the first of the continuent of self-self the below to
Arg426-Lys432	(1521)	म्युक्ताहर्ष्ट्राहर्षेत्र १५५५ (५ देहे)योग्यामुब्रोद्धर कार्य (क्रमादीर अस्टर्स)म्बर्धराम्बर्धराम्बर्धरास्त्र
Arg426-Gly431B	(1521)	कोकामन्द्रिकाचक्रमेर्याक्तामध्ये व्यक्तिमध्य । कोल्डेस्ट्रिकेन्द्रक्षमुक्तस्य । कोन्स्रिक्स्यान्य अस्
Asn425-Lys432	(1515)	रातकार । शहर करे हो के देव हैं हो हो हो हो है
Consensus	(1521)	CCTGGGCTTCCTGGGCGCCGCCGGCAGCACCATGGGCGCC
		1561 1600
Ile424-Ala433	(1549)	र व प्राप्तकार प्रयोग दिवेतर व विकास प्रदृष्टी एक हिंद विकास कार प्रदृष्टी वास हिंद विकास है।
Trp427-Gly431	(1561)	and an eight eight with mysterician in the steeming and a metally like inter-
Gln422-Tyr435B	(1537)	કે છે. આંગમાં છે. છે. જે જે તે કાર્યક્ષિયાન છે. આંગમાં માત્ર કર્યા હોય કરિયા છે. જે છે. જે જે જે જે જે જે જે જ
Arg426-Gly431	(1561)	SCORESCULLARIES NOTE OF THE OFFICE OF COLUMN OF THE CONTROL THE CO
Ile423-Met434	(1543)	्रा १ तर्गाः स्थाप्रमाणेश्वाः स्थानेत् स्थानेत् ।
Gln422-Tyr435	(1537)	where to the end of the control of the transfer of the control of
Arg426-Lys432	/1EC11	
	(1561)	mentalpolo entethetethalporenamie er del se mulejetolo deligitigetides
Arg426-Gly431B	(1561)	atrizmitota osilespa jelenji mykstraturka osiya za osilele kundiških žilenijany Kalforina (mitaan) kundali sistoni mykstratickih kundika, mitalis shebullugja njish
Arg426-Gly431B Asn425-Lys432		
	(1561)	eke kentrore, primita us infrança strusti teplije jir includeri, kultur je aflotoji noje, kili asy
Asn425-Lys432	(1561) (1555)	ere foresene formatere lationed interestativation in energial, et un francolemoje in tradi Energialistica in a serie in la storine la jarante, a provincia la la juga politikaj orga
Asn425-Lys432 Consensus Ile424-Ala433	(1561) (1555)	CGCAGCCTGACCCTGACCGTGCAGGCCCGCCAGCTGCTGA
Asn425-Lys432 Consensus	(1561) (1555) (1561)	CGCAGCCTGACCCTGACCGTGCAGGCCCGCCAGCTGCTGA
Asn425-Lys432 Consensus Ile424-Ala433	(1561) (1555) (1561) (1589)	CGCAGCCTGACCCTGACCGTGCAGGCCCGCCAGCTGCTGA  1640  1640  1640  1640  1640  1640
Asn425-Lys432 Consensus Ile424-Ala433 Trp427-Gly431	(1561) (1555) (1561) (1589) (1601)	CONTROL OF THE STATE OF T
Asn425-Lys432 Consensus Ile424-Ala433 Trp427-Gly431 Gln422-Tyr435B	(1561) (1555) (1561) (1589) (1601) (1577)	CGCAGCCTGACCCTGACCGTGCAGGCCCGCCAGCTGCTGA  1640
Asn425-Lys432 Consensus Ile424-Ala433 Trp427-Gly431 Gln422-Tyr435B Arg426-Gly431 Ile423-Met434 Gln422-Tyr435	(1561) (1555) (1561) (1589) (1601) (1577) (1601) (1583)	CGCAGCCTGACCCTGACCGTGCAGGCCCGCCAGCTGCTGA  1640
Asn425-Lys432 Consensus Ile424-Ala433 Trp427-Gly431 Gln422-Tyr435B Arg426-Gly431 Ile423-Met434 Gln422-Tyr435 Arg426-Lys432	(1561) (1555) (1561) (1589) (1601) (1577) (1601) (1583) (1577)	CGCAGCCTGACCCTGACCGTGCAGGCCCGCCAGCTGCTGA  1640
Asn425-Lys432 Consensus Ile424-Ala433 Trp427-Gly431 Gln422-Tyr435B Arg426-Gly431 Ile423-Met434 Gln422-Tyr435 Arg426-Lys432 Arg426-Gly431B	(1561) (1555) (1561) (1589) (1601) (1577) (1601) (1583) (1577) (1601) (1601)	CECAGCCTGACCCTGACCGTGCAGGCCCGCCAGCTGCTGA  1601 1640  COCAGCCTGACCCTGACCGTGCAGGCCCGCCAGCTGCTGA  1601 1640  COCAGCCTGACCCTGACCGTGCAGGCCCGCCAGCTGCTGA  1601 1640  COCAGCCTGACCCTGACCGTGCAGGCCCGCCAGCTGCTGA  COCAGCCTGACCCTGACCGTAGAGAGAGAGAGAGAGAGAGAG
Asn425-Lys432 Consensus Ile424-Ala433 Trp427-Gly431 Gln422-Tyr435B Arg426-Gly431 Ile423-Met434 Gln422-Tyr435 Arg426-Lys432	(1561) (1555) (1561) (1589) (1601) (1577) (1601) (1583) (1577) (1601) (1601)	CGCAGCCTGACCCTGACCGTGCAGGCCCGCCAGCTGCTGA  1640
Asn425-Lys432 Consensus Ile424-Ala433 Trp427-Gly431 Gln422-Tyr435B Arg426-Gly431 Ile423-Met434 Gln422-Tyr435 Arg426-Lys432 Arg426-Gly431B	(1561) (1555) (1561) (1561) (1589) (1601) (1577) (1601) (1583) (1577) (1601) (1601) (1595)	CECAGCCTGACCCTGACCGTGCAGGCCCGCCAGCTGCTGA  1601 1640  COCAGCCTGACCCTGACCGTGCAGGCCCGCCAGCTGCTGA  1601 1640  COCAGCCTGACCCTGACCGTGCAGGCCCGCCAGCTGCTGA  1601 1640  COCAGCCTGACCCTGACCGTGCAGGCCCGCCAGCTGCTGA  COCAGCCTGACCCTGACCGTAGAGAGAGAGAGAGAGAGAGAG
Asn425-Lys432 Consensus  Ile424-Ala433 Trp427-Gly431 Gln422-Tyr435B Arg426-Gly431 Ile423-Met434 Gln422-Tyr435 Arg426-Lys432 Arg426-Gly431B Asn425-Lys432 Consensus	(1561) (1555) (1561) (1561) (1589) (1601) (1577) (1601) (1583) (1577) (1601) (1601) (1595)	CGCAGCCTGACCCTGACCGTGCAGGCCCGCCAGCTGCTGA  1640
Asn425-Lys432 Consensus  Ile424-Ala433 Trp427-Gly431 Gln422-Tyr435B Arg426-Gly431 Ile423-Met434 Gln422-Tyr435 Arg426-Lys432 Arg426-Gly431B Asn425-Lys432 Consensus Ile424-Ala433	(1561) (1555) (1561) (1589) (1601) (1577) (1601) (1583) (1577) (1601) (1601) (1595) (1601)	CGCAGCCTGACCCTGACCGTGCAGGCCCGCCAGCTGCTGA  1601 1640  1630 1640  1640 1640 1640 1640 1640 1640 1640 1640
Asn425-Lys432 Consensus  Ile424-Ala433 Trp427-Gly431 Gln422-Tyr435B Arg426-Gly431 Ile423-Met434 Gln422-Tyr435 Arg426-Lys432 Arg426-Gly431B Asn425-Lys432 Consensus Ile424-Ala433 Trp427-Gly431	(1561) (1555) (1561) (1589) (1601) (1577) (1601) (1583) (1577) (1601) (1601) (1595) (1601)	CGCAGCCTGACCCTGACCGTGCAGGCCCGCCAGCTGCTGA  1601 1640  1601 1640  1601 1640  1601 1601  16
Asn425-Lys432 Consensus Ile424-Ala433 Trp427-Gly431 Gln422-Tyr435B Arg426-Gly431 Ile423-Met434 Gln422-Tyr435 Arg426-Lys432 Arg426-Gly431B Asn425-Lys432 Consensus Ile424-Ala433 Trp427-Gly431 Gln422-Tyr435B	(1561) (1555) (1561) (1589) (1601) (1577) (1601) (1583) (1577) (1601) (1601) (1601) (1629) (1641) (1617)	CGCAGCCTGACCCTGACCGTGCAGGCCCGCCAGCTGCTGA 1601 1640  1601 1640  1601 1640  1601 1600  1600  1
Asn425-Lys432 Consensus  Ile424-Ala433 Trp427-Gly431 Gln422-Tyr435B Arg426-Gly431 Ile423-Met434 Gln422-Tyr435 Arg426-Lys432 Arg426-Gly431B Asn425-Lys432 Consensus Ile424-Ala433 Trp427-Gly431	(1561) (1555) (1561) (1589) (1601) (1577) (1601) (1583) (1577) (1601) (1601) (1601) (1629) (1641) (1617)	CGCAGCCTGACCCTGACCGTGCAGGCCCGCCAGCTGCTGA  1601 1640  1630 1640  1640 1640  16

FIG. 4H

Ile423-Met434	(1623)	ericity eleteres in the position of eleteration and enteresticity in teacher of the eleteration of the electricity of the elect
Gln422-Tyr435	(1617)	अंद्राराह्यदार ,राक्त्रीक्षेत्रीयदीकार्यक्षकःभृददि संस्थेद ,यित्वाक्षेत्राधार्थः, दिश्वद्रतिसम्बद्धार्थः ।
Arg426-Lys432	(1641)	Heterologie eteratelogiaeterakete ineternie i sisterantelike telpsie ennétere i tro
Arg426-Gly431B	(1641)	interpretation interpretation and the state of the state of the state of the state of the
Asn425-Lys432	(1635)	nouncies a contelestraturas elementamentamente en entre entre elemente de la contene d
Consensus	(1641)	CGAGGCCCAGCAGCAGCTGCAGCTGACCGTGTGGGGC
		1681 1720
Ile424-Ala433	(1669)	transprogential and and projection of an end of the projection of
Trp427-Gly431	(1681)	an activation relation, electrical effectives of state of seasons of state of seasons of
Gln422-Tyr435B	(1657)	જમાન એક પ્રાપ્ત કર્યા છે. તે અને કરા કરા કર્યા કર્યા છે. તે કર્યા કર્યા કર્યા કર્યા છે છે. તે કર્યા કર્યા કર્ય
Arg426-Gly431	(1681)	equivier grand educir, egges englo de productive de principal de se principal de la company de la company de l
Ile423-Met434	(1663)	कत्रा वस्तव कृतवा ने क्वांसविक्तिक अध्याका वृद्धि होत्रे हैं व्यक्तिया है।
Gln422-Tyr435	(1657)	$u_{i}(t), v_{i}(t) \in \{c(t), t\},  a quifficite testivity of the analysis testion in the property of the pr$
Arg426-Lys432	(1681)	क्षित्राक्ष्रीयाव्यक्षका प्राथ्मकार्यवाचित्रकृत्विक्षेत्र क्षांतामक्ष्रेमचित्र व्यवस्थितिकार्यन्ति । स
Arg426-Gly431B	(1681)	कर , क्रमार , क्रमार , व्यवस्था । व्यवस्थानम् । मेर्नु द्वासीकार्यः । क्रमार । मिलाव-क्रमान्यमञ्जूकार्यः ग्रा
Asn425-Lys432	(1675)	कार्यकार माने वार्यकार में होते होते होते होते होते होते होते होते
Consensus	(1681)	ATCAAGCAGCTGCAGGCCCGCGTGCTGGAGCGCT
T1-424 31-422	/1700ì	1721 1760
Ile424-Ala433 Trp427-Gly431	(1709)	to the material action of the state of the second state of the second se
Gln422-Tyr435B	(1721)	रम ह मान हराबहर यह महिल्लाहालहा एक्ट्रांस्ट स्टाइड स्वयंत्रियार महिल्लाहा
Arg426-G1v431	(1697) (1721)	THE BOOKER CONTRACTOR STATE OF THE CONTRACTOR OF
Ile423-Met434	(1703)	pril Pritandipola Cala Milandia Robert ana 188 majar anda 24. Managana ang mangana ang Salatan
Gln422-Tyr435	(1697)	The contractions of the proposition of the contraction of the contract
Arg426-Lys432	(1721)	Yo soldmin takto citrejantiquiginekketi nismakungislatinkketiant eta
Arg426-Gly431B	(1721)	मुद्राः तसरः व्यक्तदेशकरोत्।मैहारं सम्बन्धः नद्याः (कर्णवास्तरे सम्बन्धः समानिक्तस्य । स्व
Asn425-Lys432	(1715)	the entrantial policies of the metalogical activities and caterioristic or or
Consensus	(1721)	ACCTGAAGGACCAGCAGCTGCTGGGGCATCTGGGGCTGCAG
		1761 1800
Ile424-Ala433	(1749)	ा सहित का के के सह महिल्य के लिए प्रमाणिय के के किए हैं। स्वर्ध के सिल्य के किए के किए के किए के किए के किए के
Trp427-Gly431	(1761)	Seteleta un'una especiona de Molotan lecher e Mas modeir e del camare e la fici
Gln422-Tyr435B	(1737)	interest protection, interest, in measure acception in a sourceast, in including the state of the contract of
Arg426-Gly431	(1761)	articles of the Green contribution of the articles of the contribution of the contribution of the contribution of
Ile423-Met434	(1743)	the face of the months of the finishing of the finishing of the property of the finishing o
Gln422-Tyr435	(1737)	in course interpretation control or in Figure 1988 property property of the first section and the self-self-self-self-self-self-self-self-
Arg426-Lys432	(1761)	y sime within the laws of his Bertophyle laws the splinks it employed by
Arg426-Gly431B	(1761)	भारत है। है कि के कि के कि
Asn425-Lys432	(1755)	er sejenfer statent all ner tiges all, gestichte folge fere judiciples des gestichte des verset des verfeit ein
Consensus	(1761)	CGGCAAGCTGATCTGCACCACCGCCGTGCCCTGGAACGCC
		1801 1840
Ile424-Ala433	(1789)	the constitution of the first content in the party of the constraint of the content of
Trp427-Gly431	(1801)	राष्ट्रीय रही हो। स. १८ (१८८४ वर १८८४ वर १८६४ वर १६६४) । १८६४ को १८६४ को १८५४ वर हो हो हो की हो है के दूर होती
Gln422-Tyr435B	(1777)	West the product block observables of a product school Architectural professory
Arg426-Gly431	(1801)	THE COLOURS OF THE STANDARD SHALLING CALCARD TO THE STANDARD EXPENSE LANGUAL
Ile423-Met434	(1783)	come in the translation of the control of the contr
Gln422-Tyr435 Arg426-Lys432	(1777)	under interprete deur interestation in et étaiteure : interess d'est interpreteurs set
	(1801)	are to the entire territory in the entire territories of the entire transfer and the entire transfer that the entire territories and the entire transfer that the entire territories and the entire transfer that the entire territories and the entire transfer to the entire transfer tr
Arg426-G1y431B Asn425-Lys432	(1801) (1795)	rate in Australia antica antica (confinite in Australia antica antica antica (confinite confi
Consensus	(1801)	AGCTGGAGCAACAAGAGCCTGGACCAGATCTGGAACAACA
consensus	(1001)	1841 1880
Ile424-Ala433	(1829)	the stricts, the least the class femalists special end is strictly be despected by the second
Trp427-Gly431	(1841)	Actival in manage incorrectivity or small juritable executional states in the constitution of the constitu
Gln422-Tyr435B	(1817)	Approximate steeres steerings retetributes attribute thing teacher the why we will be a second of the steering the second of the
Arg426-Gly431	(1841)	Many: ye nasilatopina asonan astrochtuleten erstemnyelerstyroasylnyetik. Ma
Ile423-Met434	(1823)	La reporting in Lighter with Leaving will eight shope high continuent merchaning service state of planties and C
Gln422-Tyr435		There is the second of the sec
	, 202. /	

FIG. 4I

) Prosticionisato principal de la facto de	(1841)	Arg426-Lys432
	(1841)	Arg426-Gly431B
	(1835)	Asn425-Lys432
) TGACCTGGATGGAGTGGGAGCGCGAGATCGACAACTACAC 1881 1920	(1841)	Consensus
attacketorijeschi enkonstilisialankeschinderinetaanatusteseastratustele. Tae	(1869)	Ile424-Ala433
	(1881)	Trp427-Gly431
<ul> <li>cognitive continuos de la propositio de la constitución d</li></ul>	(1857)	Gln422-Tyr435B
astrade je nazdpisjanigaki istracjejening rizpeja pošpi obstasjitova iz idestpriklejeniska	(1881)	Arg426-Gly431
<ul> <li>London de l'antique de la compressió de la compresentation de la completa del la completa de la completa del la completa de la completa della completa del la completa della completa della</li></ul>	(1863)	Ile423-Met434
	(1857)	Gln422-Tyr435
	(1881)	Arg426-Lys432
	(1881)	Arg426-Gly431B
	(1875)	Asn425-Lys432
1921 1960	(1881)	Consensus
	(1909)	Ile424-Ala433
	(1921)	Trp427-Gly431
	(1897)	Gln422-Tyr435B
	(1921)	Arg426-Gly431 Ile423-Met434
	(1903) (1897)	Gln422-Tyr435
	(1921)	Arg426-Lys432
	(1921)	Arg426-Gly431B
	(1915)	Asn425-Lys432
	(1921)	Consensus
1961 2000	,	
	(1949)	Ile424-Ala433
eleta, e je ga palece se en kojedinaje ne alagnanje mjeje semloveje se i kriticah elete i n	(1961)	Trp427-Gly431
संस्थाति अस्तरिः भारते अर्थक्षेत्रे गानका । "वस्तर क्रिन्तुस्य स्थान स्वीति ।	(1937)	Gln422-Tyr435B
efector (extend ) in a presupulsant as the sound in the point of the reflighering	(1961)	Arg426-Gly431
estral information with a state of the service index of the saugh extra residución	(1943)	Ile423-Met434
देश के राजित के कि के किया के राज्य के बहुत होते करते हैं। विकास किया है कि देश है जिसके स्वर्धिक के स्वर्धिक	(1937)	Gln422-Tyr435
etrici in alle, jugan in gratellongalatur. Litter y netraci spinaralistricin	(1961)	Arg426-Lys432
estream of the effect on a Consequence with one of the properties that define in	(1961)	Arg426-Gly431B
es in andresely, because with a signature conference area are solven sign	(1955)	Asn425-Lys432
GGGCCAGCCTGTGGAACTGGTTCGACATCAGCAAGTGGCT 2001 2040	(1961)	Consensus
o u, e., t. men septembre en neu ansastan, com a se contrafferet de sup	(1989)	Ile424-Ala433
के द्वार के अन्तर दिएक एक नवेश्योश्या है। विवास्तर के बड़े, का विकास संदर्भ अनुकृति अदि	(2001)	Trp427-Gly431
eur nieuwy recognistateus ny taminakta tan vateny oktopiew (zjet <mark>e)</mark> nie	(1977)	Gln422-Tyr435B
सीच र १ का एक जिल्ला क्रिक्टिकर १ हा वर्षात्र क्रिक्टिक रहार, एक जिल्ला स्वित्र हर्षात्र हार्	(2001)	Arg426-Gly431
to be a market of the set of the	(1983)	Ile423-Met434
Aller Committee of Europe teritorion de direstate de l'occident especiale de	(1977)	Gln422-Tyr435
CONTROL OF THE SECOND PROPERTY OF THE SECOND	(2001)	Arg426-Lys432
anny mengine rapa ngipapangka dan ng mana ratigah, auraha melaka 1714 dan sapat melaka 1714 dan sapat melaka melaka 1714 dan sapat mengina ngipangka pengina ngipangka ngipangka 1714 dan sapat mengina ngipangka ngipan	(2001)	Arg426-Gly431B Asn425-Lys432
GTGGTACATCAAGATCTTCATCATGATCGTGGGCGGCCTG	(1995) (2001)	Consensus
	(2001)	consensus
2041 2080 दणद्राज्यादास्त्रान्त्रान्त्रान्त्रान्त्रान्त्रान्त्रान्त्रान्त्रान्त्रान्त्रान्त्रान्त्रान्त्रान्त्रान्त्रान्त्	(2029)	Ile424-Ala433
द्यामध्यक्तरहातः अद्युद्धान्त्राच्यामध्यक्तर्भेत् । एवः द्रम्यामधः नामधाः । एक्समध्यकः प्रमाण्यक्ति। अस्		Trp427-Gly431
मुक्तर रहन्द्रद्राच्या में महास्त्रद्वद्वद्वास्त्रम् । तम्बन्धम् स्त्रा भेदास्त्र हृष्ट् । १८ । १५५० । सम्बन्ध		Gln422-Tyr435B
can elektri junijernija i aprovijela desime 144 i svizi serim nest Kaladia gelješu 114		Arg426-Gly431
experiorethe or estates the proceeding that the transfer of the state of the second	(2023)	Ile423-Met434
स्तर्भने व्यक्तिक देव विद्यमें समाप्त विश्वपृद्धको प्रशेष देव समाप्त के विस्तु रहा देव स्वकार वृद्धके हुन्।	(2017)	Gln422-Tyr435
while waters are by entry the layer entities where a market expension for the layer.	(2041)	Arg426-Lys432
ekszelefete réségütő műl as kereliképüsés szor a szolálam a milka espektiáletésésésésés	(2041)	Arg426-Gly431B

FIG. 4J

Asn425-Lys432	(2035)	દ્યાં તાલા લાદો દો દાંત તાલા દો કે
Consensus	(2041)	GTGGGCCTGCGCATCGTGTTCACCGTGCTGAGCATCGTGA
	, ,	2081 2120
Ile424-Ala433	(2069)	कृत्रीयित्वात्रम् व्यवस्थितः भूति। व्यवस्थितः भूति। व्यवस्थितः । विद्याने भूति। विद्याने ।
Trp427-Gly431	(2081)	्रवित्तिक्षित्रदेव क्षेत्रदेवक्षित्रकार्यक्ष्यात्रकार्यक्ष्यात्रक । विद्युक्तिकार्यक्षिक व्यव
Gln422-Tyr435B	(2057)	Antheter ferringet et met Ante bet eine bestehe met et eine felt eine fer eine eine eine
Arg426-Gly431	(2081)	अस्तिवास्त्रसात्राचाः स्थान्तवारा स्थान्ति अत्यावाद्यात् । स्थान्ति ।
Ile423-Met434	(2063)	here and contact their perfections of the letter contact project of the letter of the
Gln422-Tyr435	(2057)	casely is to law and close in the preference of contracts to be taken also capally in the impose size of
Arg426-Lys432	(2081)	stess procedures particular contratas de la contrata del contrata de la contrata de la contrata del contrata de la contrata del la contrata del la contrata de la contrata del la contrata de la contrata del la contrata del contrata del la contrata
Arg426-Gly431B	(2081)	Meir telwermandelejenkfetholdsiphe men pelaketenkatakeleneneit tile 1960
Asn425-Lys432	(2075)	स्थानगर्भात्रभावस्थान्त्रभावस्थानस्थानस्य । अत्र १००० वर्षास्य । भूगार्थस्य ।
Consensus	(2081)	ACCGCGTGCGCCAGGGCTACAGCCCCCTGAGCTTCCAGAC
		2121 2160
Ile424-Ala433	(2109)	erene errengerie feler eie einstehe telefert in ein teterleistliche interstehen elifietet im
Trp427-Gly431	(2121)	चौराहो । ४१४छान् अद्योद्येश्वस्थान्यस्थलको विवाद स्वत्ये (साधाराज्येसस्य को स्वाधारम्य स्वत्ये
Gln422-Tyr435B	(2097)	अभ्यक्षाक्ष्राक्ष्राक्ष्राक्ष्राच्या द्विष्यान्द्रित याचारुचाति द्विष्या (स्टार्ट्स्ट्रिस्ट्रिस्ट्रिस्ट्रिस्ट्
Arg426-Gly431	(2121)	ब्रोकादीय क्षित्रां स्टोनीस्ट्रीस्ट्रीस्ट्रीक्ट्रीक्ट्रीक्ट्रीस्ट्रीस्ट्रीस्ट्रीस्ट्रीस्ट्रीस्ट्रीस्ट्रीस्ट्री
Ile423-Met434	(2103)	अंदोल में प्रिक्त में विकास साम होते हैं के इस हो है
Gln422-Tyr435	(2097)	provide in the following of the engineers of the leader properties and the
Arg426-Lys432	(2121)	A first after services of exercises after a less after the following of the same of a
Arg426-Gly431B	(2121)	The processage on the effect of the entire of the bridge one of the entire of the enti
Asn425-Lys432	(2115)	રાજ્ય રજ્યાન મહી અને દ્રાંત માત્ર જિલ્લો ફેલ્સ રહેલા કરાય માત્ર માન્ય માત્ર મુખ્યાના જાણ જાણ જાણ માત્ર માત્ર મ
Consensus	(2121)	CCGCTTCCCCGCCCCCGCGCCCCGACGGC
		2161 2200
Ile424-Ala433	(2149)	કુંકાલુંક હતું.તાલાકારાન ફિલ્ફારેલ બિલી નુકારણ કરો હતું કરો કરો કરો કહે કરો છે. જે બીલી <mark>એટ કુંદરે કરે ક</mark>
Trp427-Gly431	(2161)	्रमातः । सन्तरमञ्ज्ञानन्त्रियोद्दाद्दिञ्ज्ञ दाद्देश्यक्षक्षम् । सन्दर्भन् । सन्दर्भन् । सन्दर्भन् । सन्दर्भन् । इति
Gln422-Tyr435B	(2137)	क्षर्भाव्यक्तिक्षित्रम् विभिन्नित्रिक्ष्यं विभिन्नित्र विभिन्नित्र विभिन्नित्र विभिन्नित्र विभिन्नित्र विभिन्न
Arg426-Gly431	(2161)	कृत्रा चार्या भारतिसेवबुव संबद्धायम्याति । भारतस्यवाद्यायास्य
Ile423-Met434	(2143)	१५०५ (र प्रमा केंग्रदानम्म। विद्यालेखाने स्थाति । क्षेत्रका । विद्यालेखान
Gln422-Tyr435	(2137)	त्रमा क्षेत्रकार मानव देवाकी दावानावाय हो कर्तिक है। विक्रिक्ट दावाना हेवाकी कार्य प्रकार होने विक्रिक्त कर व
Arg426-Lys432	(2161)	go no re reformación againstach talon con tach i confidilide a cal little con
Arg426-Gly431B	(2161)	कर्मार है कर है। क्षेत्रक केंग्रेस केंग्रिक स्वेदीई कार्यन्त्रक कर कार्यन क्षेत्रक है। कार्यन्त्रिक है।
Asn425-Lys432	(2155)	क विभिन्ने क्षेत्र क्षेत्र क्षेत्र क्षेत्र के विभिन्न क्षेत्र क्षेत्र क्षेत्र क्षेत्र क्षेत्र क्षेत्र के विभिन्न क
Consensus	(2161)	ATCGAGGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG
<b>73</b> 404 54 455		2201 2240
Ile424-Ala433	(2189)	Education and the state of the
Trp427-Gly431	(2201)	्रहत पर प्रारम्भव्यक्तिमा राज्यक्तात् व कर्नुन्तर राज्य राज्य क्रिक्ट प्रवस्तिकार ।
Gln422-Tyr435B	(2177)	the factor and starting of the property of the factors of the fact
Arg426-Gly431	(2201)	ment him migrationed entry place most systems in your space, the gradient additional
Ile423-Met434	(2183)	ा । ४० . प्रारम्भ विकासी स्ट्रेसिस स्ट्रिक्ट स्ट्रेसिस स्ट्रिक्ट स्ट्रेसिस स्ट्रेसिस स्ट्रेसिस स्ट्रेसिस स्ट्र
Gln422-Tyr435	(2177)	state of the more individual states and indicated and individual individual states and i
Arg426-Lys432	(2201)	त्या । योक्ता व्यवस्थित के देवी योक्ता के द्वारा के विकास स्थापन करते हैं के विकास विकास के प्रतिस्था है है ।
Arg426-Gly431B	(2201)	and the state of t
Asn425-Lys432	(2195)	The second distriction and provided the second distriction of the seco
Consensus	(2201)	GCCCCCTGGTGCACGGCCTGCTGGCCCTGATCTGGGACGA
Ile424-Ala433	(2220)	2241 2280
Trp427-Gly431	(2229)	प्रकारमान् । इ.स.चे.) एतुर्वे के मुक्त से दर्श का के का कि का कि का कि का कि का कि
Gln422-Tyr435B	(2241)	ૡ૽૽૽ૢ૱ૡ૽ૹૢ૽ૺૡ૽૽૱ૡ૽ૡ૽ૡ૽ૡૡ૽ૡૡૢૡૡૡ૽૽૱ૡૡ૽૽ઌૡૡૡ૽ૡૡૡૡૡ૽ૢૡ૽ૢૡ <b>ૡ</b> ૾ૡ <u>ૺ</u>
Arg426-Gly431	(2217)	eje njejejeleti rijete jepitovijej indigete sje o rijejo indike jestijeje lidicija jajete rijej
Ile423-Met434	(2241) (2223)	egangriculise bactora tilogistica (p. p. p. p. 170) e tilogis protectife i c. p. lonnegesett
Gln422-Tyr435	(2223) $(2217)$	go desilas suglicadas detatam clima grava, requisiria ele arrigitar sec ete visco atas tractalmente atraca como con consecución.
Arg426-Lys432		त्रीतः प्रोत्तरः प्रतिकृतिकार्यक्षणार्थव्यविद्वाति प्राप्तिकारः । तथा स्वतिकार्यक्षणार्थाः । विद्वाराध्यः । वि विकास
Arg426-Gly431B	(2241) (2241)	elentes er en er elette blevelet i en er en en en en en en en elette en ken i seg elette er en en er elette beween en
Asn425-Lys432	(2235)	edici of quaesy contest (pringigeness) i officiale en force, printaga, le baggio totul e ontago e vieg. Tra transesta e rebiete e e monte la dicinación e o como il sego for teste e totale e entre la vieg.
Consensus		CCTGCGCAGCCTGTGCCTGTTCAGCTACCACCGCCTGCGC
-4.10611343	(CC41)	
		FIG. 4K

		2281	2320
Ile424-Ala433	(2269)	untretenthagtangelängkannähteteletotaneneteltangetangen til	
Trp427-Gly431	(2281)		155 B 55
Gln422-Tyr435B	(2257)		
Arg426-Gly431	(2281)		14.535.6
Ile423-Met434	(2263)	Before maliginalism by a commercial state with the balance work.	0.00000000
Gln422-Tyr435	(2257)	दोर्गकार नहरूरोत्रां अंगोद्धियाः होहि। बीटाम योकस्वरोर को विमित्रिक्सिस्ह ।	(equivere
Arg426-Lys432	(2281)	त्रीत्री । विभागतीयां विभागतां प्रकारतां विदेश के अन्य विदेश पूर्व होते हा पूर्व होते हैं।	ide iver
Arg426-Gly431B	(2281)		Series
Asn425-Lys432	(2275)	र र र इस्टें कर इस ए स्था को काल हु इस है कि इस हो आहा है। ए ए ए ए ए ए ए ए ए ए ए ए ए ए ए ए ए ए ए	(0)0/4/0/0
Consensus	(2281)	GACCTGATCCTGATCGCCGCCCGCATCGTGGAGCT	
		2321	2360
Ile424-Ala433	(2309)	ute intide te min tegene attelminete, de la se en canalistit e prijetiele fe	J. 1. 33.4.
Trp427-Gly431	(2321)	educificacipante papilos desele prodesión e paestraly. A sultra de telefoses	25. Ker . 323
Gln422-Tyr435B	(2297)	sport in the complete party party party and the property of a second party of the complete control of the complete control of the control of	" p: (c ' 5 ' j)
Arg426-Gly431	(2321)	ender feet fit interesting in interests in a next hypometric prefetable to	12/23/6 1/12/
Ile423-Met434	(2303)	an interest and the second particles of the second second and the second second	
Gln422-Tyr435	(2297)	Arte petalogicio, cretochétora pictore colour, provione pophetoletific	
Arg426-Lys432	(2321)	ansociatore provide and this protocol postacion have been imate	
Arg426-Gly431B	(2321)	dejudicis us injulo limist piratiqualosis, umsis essajostalisis seg	
Asn425-Lys432	(2315)	्र विकास के त्रीय जिल्हा का बार देश देश हों हो । अन्य अंद राष्ट्र है । अन्य के स्वार्थ के अवदेश का अदेश हैं है	strike out
Consensus	(2321)	GCCGCCGCGCTGGGAGGCCCTGAAGTACTGGGGC	
		2361	2400
Ile424-Ala433	(2349)	an income in indicion collegation desired grant	E 0 5
Trp427-Gly431	(2361)	स्ति । १ र १५ द्राप्त के मान्यमा साहा प्राप्त । १ र १६ १६ १५ द्राप्त । विकास सम्बद्धां स्वतंत्र स्वतंत्र स्वतं	क्ष्युद्ध, जार्जा
Gln422-Tyr435B	(2337)	त्यार अभूतिकातिक कार्यक राष्ट्रांद्यां स्थापन हार्यक्षा त्रिकेट विश्वास्था स्थापन हें विश्वास्था स्थापन हें व	wer in
Arg426-Gly431	(2361)	ed a skatalistica ti ka a jedel talija jedine teoritege – njedini jedini	cials a
Ile423-Met434	(2343)	also high control water in the trip of a political reference in the sample of the first in-	6,0 0 0,5
Gln422-Tyr435	(2337)	et. It i netermentario i tam i descriptioni to compresentario per e	ଅନ୍ତମ ଓ
Arg426-Lys432	(2361)	क्षरप्रभव । क्षांकर करों । भू व्यक्ति । १,० हिम्मी हेम्बू कृष्ट्य हैं। एको विशेषा कर्त्र प्रमृत्य क्ष	වැන්දෙර
Arg426-Gly431B	(2361)	at we are far that was considered by an about a company of a reference of the	रहरत्वधारीको ह
Asn425-Lys432	(2355)	the set accounting to contract to the think on a cost of the effect of	
Consensus	(2361)	GCTGCAGTACTGGATCCAGGAGCTGAAGAACAGCG	CCGTG
		2401	2440
Ile424-Ala433	(2389)	regard seem of the Mark Commence of the Commen	e er e
Trp427-Gly431	(2401)	MERCONDINAL SECTION OF THE CONTRACTOR OF THE SECTION OF THE SECTIO	$(AG_j^{\dagger}U_{ij}), \forall j$
Gln422-Tyr435B	(2377)	man in monesterphysical physical color in the interestication confi	Section 1
Arg426-Gly431	(2401)	province of the interference of the property of the contraction of the property of the contraction of the co	gur (15
Ile423-Met434	(2383)	प्रदेशक र ्रांतर्वात गाँव (दार्विक प्रदेश कर्मुम्पाद वित्त कर्मात्रात कर्मा व्यवस्थित वास्ति (वास्ति वास्ति वास्	
Gln422-Tyr435	(2377)	the little and experience of extractions in production relative souther patients in the leave of the section of	3231130
Arg426-Lys432	(2401)	. इत्तर । । । । । । । । । । । । । । । । । । ।	
Arg426-Gly431B	(2401)	to an extension of transfer to the bit of planers by classifications.	त्रदा <u>श्</u> रूष
Asn425-Lys432	(2395)	the improvement, quantity, provincing the following places of principal energy provincing principal energy and in the following principal energy provincing provincin	
Consensus	(2401)	AGCCTGTTCGACGCCATCGCCGTGGCCGAC	3GGCA
T10424 B1-400	10100:	2441	2480
Ile424-Ala433	(2429)	and the control of th	Span.
Trp427-Gly431	(2441)	हर क्या मान्त्रिकेट प्रमान्त्रिकोड लोक मन्त्रिकोड होत्याच्या प्रमानिकोड प्रमान होते हिन्	re, cleye
Gln422-Tyr435B	(2417)	re province that a proposition tended to be relieved the left of the left of	luji ras
Arg426-Gly431	(2441)	e a en priesta i para Anto Esperaficade (a la gerrade prifotesa da	(로): Refe
Ile423-Met434	(2423)	e se production particles in the second section of the section of the second section of the section of the second section of the second section of the sectio	
Gln422-Tyr435	(2417)	्रा त्याराम् १९ व त्यात्राम् अस्ति । त्यार्थानीमा महाद्वारीत् ( १८४४) अ <b>स्ति ।</b> स्वति ।	
Arg426-Lys432	(2441)	in partie et austrescondere nederak anterestablisch bei eine	
Arg426-Gly431B	(2441)	20, an in activities in states relative and interestinations	
Asn425-Lys432	(2435)	त्यस्याम् देवः व्यक्षित्रस्याः द्वीयां द्वाराम् विक्रियं । द्वित्यं विक्रियं हर्षा स्वतः ।	(सहार्यः
Consensus	(2441)	CCGACCGCATCATCGAGGTGGCCCAGCGCATCGGCC	
T1 6424 . N1 = 422	124621	2481	2520
Ile424-Ala433	(2469)	economic legals, a creat tractical filtra in List and Alfred Alfred a tractical states and	03500

FIG. 4L

Trp427-Gly431	(2481)	्रामानद्वाद्व-दि <b>ल्देरी</b> वारमात्वादर्दाः (दिवादोऽ।दोर-प्रेन्न्यदेशस्य कारावादादाः असर्वादिकार्
Gln422-Tyr435B	(2457)	
Arg426-Gly431	(2481)	
Ile423-Met434	(2463)	
Gln422-Tyr435	(2457)	
Arg426-Lys432	(2481)	strong fee felestrongram explore monderplanes administrately in the conte
Arg426-Gly431B	(2481)	कार्यका सम्बद्धिक एक प्रवेश के बोर्च के अपने के प्रवेश के प्रवेश के प्रवेश के कि एक प्रवेश के प्रवेश के प्रवेश के
Asn425-Lys432	(2475)	controller te maneral operative and a le son establishe logishe de logishe territorie substantion of
Consensus	(2481)	CTTCCTGCACATCCCCCGCCGCATCCGCCAGGGCTTCGAG
		2521 2541
Ile424-Ala433	(2509)	edelo lette folkledelometrik oldstations, like de
Trp427-Gly431	(2521)	Mark that have received electrophysics on a constr-
Gln422-Tyr435B	(2497)	especial britarians and the management of the state of th
Arg426-Gly431	(2521)	अवस्था को हो रहा है। में करान पूर्व साम के स्थापन है। जिल्ला
Ile423-Met434	(2503)	cheforetise (e) teers collects water to collect.
Gln422-Tyr435	(2497)	with the Conference of the State of the Stat
Arg426-Lys432	(2521)	cheks ministering a malaysing sykelsüla jarena
Arg426-Gly431B	(2521)	्राकृति द्वार । वर्षकृत्र वेत्राव्य वर्षाकृतिकृतिकृतिक वर्षाव । वर्षाक
Asn425-Lys432	(2515)	MOTO CONTROL BANDO DE DA PARA DE LA PARA DEL PARA DE LA
Consensus	(2521)	CGCGCCCTGCTGTAACTCGAG

FIG. 4M

110 00/3/303	28	/ <sub>1</sub> 65
		-
Leu122-Ser199-Tryp427-G1y431	(1)	GAATTCGCCACCATGGATGCAATGAAGAGA
Val127-Asn195-Arg426-Gly431	(1)	GAATTEGECACCATGGATGCAATGAAGAGA
Val120-Thr202-Ile424-Ala433	(1)	GAATTCGCCACCATGGATGCAATGAAGAGA
Leu122-Ser199-Arg426-Lys432	(1)	GAATTCGCCACCATGGATGCAATGAAGAGA
Leu122-Ser199-Arg426-Gly431	(1)	GAATTCGCCACCATGGATGCAATGAAGAGA
Lys121-Val200-Asn425-Lys432	(1)	GAATTCGCCACCATGGATGCAATGAAGAGA
Val120-Ile201-Ile424-Ala433	(1)	GAATTCGCCACCATGGATGCAATGAAGAGA
Vall20-Ile201B-Ile424-Ala433	(1)	GAATTČGCCACCATGGATGCAATGAAGAGA
Consensus	(1)	GAATTCGCCACCATGGATGCAATGAAGAGA
		31 60
Leu122-Ser199-Tryp427-Gly431	(31)	GGGCTCTGCTGTGTGCTGCTGTGGA
Val127-Asn195-Arg426-Gly431	(31)	GGGCTCTGCTGTGTGCTGCTGTGTGGA
Val120-Thr202-Ile424-Ala433	(31)	GGGCTCTGCTGTGTGCGA
Leu122-Ser199-Arg426-Lys432	(31)	GGGCTCTGCTGTGTGGA
Leu122-Ser199-Arg426-Gly431	(31)	GGGCTCTGCTGTGTGCA
Lys121-Val200-Asn425-Lys432	(31)	GGGCTCTGCTGTGTGCTGCTGTGGA
Val120-Ile201-Ile424-Ala433	(31)	GGGCTCTGCTGTGTGCA
Val120-Ile201B-Ile424-Ala433	(31)	GGGCTCTGCTGTGTGCA
Consensus	(31)	GGGCTCTGCTGTGTGCTGCTGTGTGGA
T122 0100 m		61 90
Leu122-Ser199-Tryp427-Gly431	(61)	GCAGTCTTCGTTTCGCCCAGCGCCGTGGAG
Vall27-Asn195-Arg426-Gly431	(61)	GCAGTCTTCGTTTCGCCCAGCGCCGTGGAG
Vall20-Thr202-Ile424-Ala433	(61)	GCAGTGTTCGTTTCGCCCAGCGCCGTGGAG
Leu122-Ser199-Arg426-Lys432	(61)	GCAGTCTTCGTTTCGCCCAGCGCCGTGGAG
Leu122-Ser199-Arg426-Gly431	(61)	GCAGTGTTCGTTTCGCCCAGCGCCGTGGAG
Lys121-Val200-Asn425-Lys432	(61)	GCAGTCTTCGTTTCGCCCAGCGCCGTGGAG
Vall20-Ile201-Ile424-Ala433	(61)	GCAGTCTTCGTTTCGCCCAGCGCCGTGGAG
Vall20-Ile201B-Ile424-Ala433	(61)	GCAGTCTTCGTTTCGCCCAGCGCCGTGGAG
Consensus	(61)	GCAGTCTTCGTTTCGCCCAGCGCCGTGGAG
I 122	(01)	91 120
Leu122-Ser199-Tryp427-Gly431	(91)	AAGCTGTGGGTGACCGTGTACTACGGCGTG
Vall27-Asn195-Arg426-Gly431	(91)	AAGCTGTGGGTGACCGTGTACTACGGCGTG
Val120-Thr202-Ile424-Ala433	(91)	AAGCTGTGGGTGACCGTGTACTACGGCGTG
Leu122-Ser199-Arg426-Lys432	(91)	AAGCTETGGGTGACCGTGTACTACGGCGTG
Leu122-Ser199-Arg426-Gly431	(91)	AAGCTGTGGGTGACCGTGTACTAGGGCGTG
Lys121-Val200-Asn425-Lys432	(91)	AAGGTETGEGTGACCGTGTACTAGGGGGTG
Val120-Ile201-Ile424-Ala433	(91)	AAGCTGTGGGTGACCGTGTACTAGGGGGTG
Val120-Ile201B-Ile424-Ala433	(91)	AAGCTGTGGGTGACCGTGTACTACGGCGTG
Consensus	(91)	AAGCTGTGGGTGACCGTGTACTACGGCGTG
Y100 C. 100 T. 407 C. 421	(101)	121 150
Leu122-Ser199-Tryp427-Gly431	(121)	CCCGTGTGCAAGGAGGCCACCACCCTG
Val127-Asn195-Arg426-Gly431	(121)	CCGGTGTGGAAGGAGGCCACCACCACCGTG
Val120-Thr202-Ile424-Ala433	(121)	CCCGTCTGCAAGGAGGCCACCACCACCTG
Leu122-Ser199-Arg426-Lys432	(121)	PROFILE TO A SECULAR CONTRACTOR OF THE PROFILE OF T
Leu122-Ser199-Arg426-Gly431	(121)	CCCCCGGGGGGGGGGCCACCACCGCGG
Lys121-Val200-Asn425-Lys432	(121)	CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
Val120-Ile201-Ile424-Ala433		CCCGTGTGGAAGGAGGCCACCACCACCTG
Val120-Ile201B-Ile424-Ala433	(121)	and the second s
Consensus	(121)	CCCGTGTGGAAGGAGGCCACCACCACCCTG
Tau 100 0 100 m 100 - 51 100		151 180
Leu122-Ser199-Tryp427-Gly431	(151)	TTCTGCGCCAGCGACGCCAAGGCCTACGAC
Vall27-Asn195-Arg426-Gly431	(151)	TTCTGCGCCAGCGCCAAGGCCTACGAC
Val120-Thr202-Ile424-Ala433	(151)	TTETGCGCCAGCGACGCCAAGGCCTACGAC
Leu122-Ser199-Arg426-Lys432	(151)	TTCTGEGCCAGCGACGCCAAGGCCTACGAC
Leu122-Ser199-Arg426-Gly431	(151)	TTCTGCGCCAGGGCCAAGGCCTACGAC
Lys121-Va1200-Asn425-Lys432	(151)	TTCTGCGCCAGCGACGCCAAGGCCTACGAC

WO 00/39303

PCT/US99/31272

WO 00/39303	29	/ 65	PCT/US99/31272
Val120-Ile201-Ile424-Ala433	(151)	TTCTGCGCCAGCGAC	CCCANCCCCMNCCC
Val120-Ile201B-Ile424-Ala433	(151)	TTCTGCGCCAGCGAC	
Consensus	(151)		
Consensus	(131)	181	210
Leu122-Ser199-Tryp427-Gly431	(181)		
Val127-Asn195-Arg426-Gly431	(181)	ACCGAGGTGCACAAC	GTGTGGGCCACCCAC
Val120-Thr202-Ile424-Ala433	(181)		GTGTGGGCCACCCAC
Leu122-Ser199-Arg426-Lys432	(181)	ACCGAGGTGCACAAC	
Leu122-Ser199-Arg426-Gly431	(181)	ACCGAGGTGCACAAC	
Lys121-Val200-Asn425-Lys432	(181)		
Val120-Ile201-Ile424-Ala433	(181)	ACCGAGGTGCACAAC	
Vall20-Ile201B-Ile424-Ala433	(181)	ACCGAGGTGCACAAC	
Consensus	(181)		
00.100.1000	(101)	211	240
Leu122-Ser199-Tryp427-Gly431	(211)	<del>-</del>	
Val127-Asn195-Arg426-Gly431	(211)		
Val120-Thr202-Ile424-Ala433	(211)	a transfer of the property of the second	
Leu122-Ser199-Arg426-Lys432	(211)		
Leu122-Ser199-Arg426-Gly431	(211)	A STATE OF THE PROPERTY OF THE PARTY OF THE	
Lys121-Val200-Asn425-Lys432	(211)	GCCTGCGTGCCCACC	
Val120-Ile201-Ile424-Ala433		GCCTGCGTGCCCACC	SACCCCAACCCCCAC
Val120-Ile201B-Ile424-Ala433	(211)	GCCTGCGTGCCCACC	
Consensus		GCCTGCGTGCCCACC	
	,,	241	270
Leu122-Ser199-Tryp427-Gly431	(241)	GAGATCGTCCTGGAG	
Val127-Asn195-Arg426-Gly431	(241)	GAGATEGTGCTGGAG	AACGTGACCGAGAAC
Val120-Thr202-Ile424-Ala433	(241)	GAGATCGTGCTGGAG	
Leu122-Ser199-Arg426-Lys432	(241)	GAGATCGTGCTGGAG	
Leu122-Ser199-Arg426-Gly431	(241)	GAGATCGTGCTGGAGA	
Lys121-Va1200-Asn425-Lys432	(241)	GAGATCGTGCTGGAG	
Val120-Ile201-Ile424-Ala433	(241)	GAGATCGTGCTGGAG	
Val120-Ile201B-Ile424-Ala433	(241)	GAGATCGTGCTGGAG/	AACGTGACCGAGAAC
Consensus	(241)	GAGATCGTGCTGGAGA	
		271	300
Leu122-Ser199-Tryp427-Gly431	(271)	TTCAACATGTGGAAGA	ACAACATGGTGGAG
Val127-Asn195-Arg426-Gly431	(271)	TTCAACATGTGGAAGA	
Val120-Thr202-Ile424-Ala433	(271)	DIVERSOR VEHICLES	
Leu122-Ser199-Arg426-Lys432	(271)	HIVE A CONTENED AND	
Leu122-Ser199-Arg426-Gly431	(271)	TTCAACATGTGGAAGA	
Lys121-Val200-Asn425-Lys432	(271)	TTCAACATGTGGAAGA	
Val120-Ile201-Ile424-Ala433	(271)	TTCAACATGTGGAAGA	
Val120-Ile201B-Ile424-Ala433	(271)	TTCAACATGTGGAAGA	ACAACATGGTGGAG
Consensus	(271)	TTCAACATGTGGAAGA	ACAACATGGTGGAG
		301	330
Leu122-Ser199-Tryp427-Gly431	(301)	<b>EAGATGCACGAGGACA</b>	TCATICAGE GUEUGG
Val127-Asn195-Arg426-Gly431	(301)	CAGATGCACGAGGACA	
Val120-Thr202-Ile424-Ala433	(301)	CAGATGCACGAGGACA	TCATCAGCCTGFGG
Leu122-Ser199-Arg426-Lys432	(301)	CAGATGCACGAGGACA	TCATCAGCCTGTGG
Leu122-Ser199-Arg426-Gly431		CAGATGCACGACGACA	
Lys121-Va1200-Asn425-Lys432		CAGATGCACGAGGACA	
Val120-Ile201-Ile424-Ala433		CAGATGCACGAGGACA	
Vall20-Ile201B-Ile424-Ala433	(301)	CAGATGCACGAGGACA	TCATCAGCCTGTGG
Consensus	(301)	CAGATGCACGAGGACA	TCATCAGCCTGTGG
		331	360
Leul22-Ser199-Tryp427-Gly431		GACCAGAGCCTGAAGC	
Val127-Asn195-Arg426-Gly431	(331)	GACCAGAGCCTGAAGC	CCTGCGTGAAGCTG
Val120-Thr202-Ile424-Ala433	(331)	GACCAGAGCCTGAAGC	CCTGCGTG

WO 00/39303	30	/ 65 PCT/US99/31272
Leu122-Ser199-Arg426-Lys432	(331)	GACCAGAGCCTGAAGCCCTGCGTGAAGCTG
Leu122-Ser199-Arg426-Gly431	(331)	GACCAGAGCCTGAAGCCCTGCGTGAAGCTG
Lys121-Val200-Asn425-Lys432	(331)	GACCAGAGCCTGAAGCCCTGCGTGAA
Val120-Ile201-Ile424-Ala433	(331)	GACCAGAGCCTGAAGCCCTGCGTG
Val120-Ile201B-Ile424-Ala433	(331)	GACCAGAGCCTGAAGCCCTGCGTG
Consensus	(331)	GACCAGAGCCTGAAGCCCTGCGTGAAGCTG
		361 390
Leul22-Ser199-Tryp427-Gly431	(361)	
Val127-Asn195-Arg426-Gly431	(361)	ACCCCCTGTGCGTGGGGGCAGGGAACTGC
Val120-Thr202-Ile424-Ala433	(355)	
Leu122-Ser199-Arg426-Lys432	(361)	GG
Leu122-Ser199-Arg426-Gly431	(361)	GG
Lys121-Val200-Asn425-Lys432	(357)	GG
Val120-Ile201-Ile424-Ala433	(355)	
Val120-Ile201B-Ile424-Ala433	(355)	
Consensus	(361)	GG
	,	391 420
Leul22-Ser199-Tryp427-Gly431	(363)	CAACAGCGTGATCACCCAGGCCTGCCCC
Val127-Asn195-Arg426-Gly431	(391)	AACACCAGCGTGATCACCCAGGCCTGCCCC
Val120-Thr202-Ile424-Ala433	(357)	CGGCGCCACCCAGGCCTGCCCC
Leu122-Ser199-Arg426-Lys432	(363)	CAACAGCGTGATCACCCAGGCCTGCCCC
Leu122-Ser199-Arg426-Gly431	(363)	CAACACCGTGATCACCCAGGCCTGCCCC
Lys121-Val200-Asn425-Lys432	(359)	CECCEGEGATCACCCAGGCCTGCCCC
Val120-Ile201-Ile424-Ala433		GEGGCATCACCCAGGCCTGGCCC
Vall20-Ile201B-Ile424-Ala433	(355)	CCCGGCATCACCCAGGCCTGCCCC
Consensus	(391)	CA CAGCGTGATCACCCAGGCCTGCCCC
	,,	421 450
Leul22-Serl99-Tryp427-Gly431	(391)	AAGGTGAGCTTCGAGCCCATCCCCATCCAC
Val127-Asn195-Arg426-Gly431	(421)	AAGGTEAGETT CGACCCCATCCCCATCCAC
Val120-Thr202-Ile424-Ala433	(379)	AAGGTGAGCTECGAGCCCATCCCATCCAC
Leu122-Ser199-Arg426-Lys432	(391)	AAGGTEAGCTTEGAGCCCATCCCCATCCAC
Leu122-Ser199-Arg426-Gly431	(391)	AAGGTGAGGTTEGAGCCCATGCCGATGGAG
Lys121-Val200-Asn425-Lys432	(385)	AAGGTGAGCTTCGAGCCCATCCCATCCAC
Val120-Ile201-Ile424-Ala433	(379)	AAGGIERGE VEGAGGCCATICCCCATICCAC
Vall20-Ile201B-Ile424-Ala433	(379)	AAGGHEAG
Consensus	(421)	AAGGTGAGCTTCGAGCCCATCCCCATCCAC
0000040	, ,	451 480
Leu122-Ser199-Tryp427-Gly431	(421)	PANOTICISCO SOCCESCO GENERAL GEO ALCOTG
Val127-Asn195-Arg426-Gly431	(451)	wilestes costs acceded to wide CC: wise TG
Val120-Thr202-Ile424-Ala433	(409)	##/@stellectors/secondedenhiledede/hiseeffG
Leu122-Ser199-Arq426-Lys432	(421)	INTERCOLOGICA CONTROL
Leu122-Ser199-Arg426-Gly431	(421)	With Control of the c
Lys121-Val200-Asn425-Lys432	(415)	лу, се, Сессе се
Val120-Ile201-Ile424-Ala433	(409)	୲୰ଽୡୢ୷୶୶୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷୷
Val120-Ile201B-Ile424-Ala433	(409)	W/Cileseccesseccescomyecces/vice/iG
Consensus	(451)	TACTGCGCCCCGCCGGCTTCGCCATCCTG
Conscrisus	(201)	481 510
Leu122-Ser199-Tryp427-Gly431	(451)	AAGTGGAAGGAGAAGTTCAACGGCAGC
Val127-Asn195-Arg426-Gly431		AAGTGCAAGACAAGAAGTTCAACGGCAGC
Vall20-Thr202-Ile424-Ala433	(439)	AAGTGCAAGGACAAGAAGTTCAACGGCAGC
Leu122-Ser199-Arg426-Lys432	(451)	ANGTGEANCEACAGAAGTTCAACGGCAGC
Leu122-Ser199-Arg426-Gly431		AAGTGCAACAAGAAGTTCAACGGCAGC
Lys121-Val200-Asn425-Lys432		AAGTGCAACGACAAGAAGTTCAACGGCAGC
Val120-Ile201-Ile424-Ala433		AAGTGCAACGACAAGAAGTTCAACGGCAGC
Val120-11e201-11e424-Ala433	14301	AAGTGTAACGACAAGAAGTTCAACGGCAGC
Consensus	(401)	AAGTGCAACGACAAGAAGTTCAACGGCAGC
Consensus	(401)	511 540
		217

WO 00/39303	21	1 65	DCT#1500/31272
	. <b>31</b> (481)	/ 65	PCT/US99/31272
Leu122-Ser199-Tryp427-Gly431 Val127-Asn195-Arg426-Gly431	(511)	GGCCCTGCACCAACGTC	
Val120-Thr202-Ile424-Ala433	(469)	GGCCCTGCACCAACGT	
Leu122-Ser199-Arg426-Lys432	(481)	GGCCCCTGCACCAACGTG	
Leu122-Ser199-Arg426-Gly431	(481)	GGCCCCTGCACCAACGT	
Lys121-Val200-Asn425-Lys432	(475)	GGCCCCTGCACCAACGTC	
Val120-Ile201-Ile424-Ala433	(469)	GGCCCCTGCACCAACGTC	Chart the Park and Car and Control of the Control o
Val120-Ile201B-Ile424-Ala433	(469)	GGCCCCTGCACCAACGTG	
Consensus	(511)	GGCCCTGCACCAACGTG	
		541	570
Leu122-Ser199-Tryp427-Gly431	(511)	TGCACCCACGCCATCCGC	CCCGTGGTGAGC
Val127-Asn195-Arg426-Gly431	(541)	TGCACCCACGCCATCCGC	
Val120-Thr202-Ile424-Ala433	(499)	TGCACCCACGGCATCCGC	
Leu122-Ser199-Arg426-Lys432	(511)	TGCACCCACGGCATCCGC	
Leu122-Ser199-Arg426-Gly431	(511)	TGCACCCACGGCATCCGC	
Lys121-Val200-Asn425-Lys432	(505)	TGCACCCACGGCATCCGC	
Val120-Ile201-Ile424-Ala433	(499)	TGCACCCACGGCATCCGC	
Vall20-Ile201B-Ile424-Ala433	(499)	TGCACCCACGGCATCCGC	and the state of t
Consensus	(541)	TGCACCCACGGCATCCGC	
Leu122-Ser199-Tryp427-Gly431	(541)	571 ACCCAGCTGCTGCTGAAC	600
Vall27-Asn195-Arg426-Gly431	(571)	ACCCAGCTGGTGCTGAAC	
Val120-Thr202-Ile424-Ala433	(529)	ACCEAGCEGETGAAC	
Leu122-Ser199-Arg426-Lys432	(541)	ACCCAGCTCCTGCTGAAC	
Leu122-Ser199-Arg426-Gly431	(541)	ACCCAGCTGCTGCTGAAC	
Lys121-Val200-Asn425-Lys432	(535)	ACCCAGCTGCTGCTGAAC	
Val120-Ile201-Ile424-Ala433	(529)	ACCCAGCTGCTGCTGAAC	
Val120-Ile201B-Ile424-Ala433	(529)	ACCCAGCTGCTGCTGAAC	
Consensus	(571)	ACCCAGCTGCTGCTGAAC	GGCAGCCTGGCC
		601	630
Leul22-Ser199-Tryp427-Gly431	(571)	GAGGAGGCGTGGTGATC	
Val127-Asn195-Arg426-Gly431	(601)	GAGGAGGGCGTGGTGATC	
Val120-Thr202-Ile424-Ala433	(559)	GAGGAGGCCTGGTGATC	
Leu122-Ser199-Arg426-Lys432	(571)	<b>सः</b> यसः ग्रह्मयव्यक्तम्बस्यास्य ।	
Leul22-Ser199-Arg426-Gly431 Lys121-Val200-Asn425-Lys432	(571)	GAGGAGGGCGTGGTGATC	CGCAGCGAGAAC
Vall20-Ile201-Ile424-Ala433	(565)	GAGGAGGCGTGGTGATC	
Val120-11e201-11e424-Ala433	(559) (559)	GAGGAGGGGGTGGTGATE	
Consensus	(601)	GAGGAGGGCGTGGTGATC	
consensus	(001)	631	660
Leu122-Ser199-Tryp427-Gly431	(601)	TTCACCGACAACGCCAAG	
Val127-Asn195-Arg426-Gly431	(631)	TTCACCGACAACGCCAAG	
Val120-Thr202-Ile424-Ala433	(589)	TTCACCGACAACGCCAAG	the analysis of the second contract of the se
Leu122-Ser199-Arg426-Lys432	(601)	TTEACCGACAACGCCAAG	
Leu122-Ser199-Arg426-Gly431	(601)	TICACOGACAACGCCAAG	
Lys121-Va1200-Asn425-Lys432	(595)	TUCACCGACAACGCCAAG	ACCATCATUGEG
Val120-Ile201-Ile424-Ala433		TICACEGACAACGCCAAG	
Vall20-Ile201B-Ile424-Ala433		TICACCGACAACGCCAAG	
Consensus	(631)	TTCACCGACAACGCCAAG	
1100 0 100 =		661	690
Leu122-Ser199-Tryp427-G1y431		CAGCTGAAGGAGAGCGTG	
Vall27-Asn195-Arg426-Gly431		CAGCTGAAGGAGAGCGTG	
Val120-Thr202-Ile424-Ala433 Leu122-Ser199-Arg426-Lys432		CAGCTGAAGGAGAGCGTG	
Leu122-Ser199-Arg426-Lys432 Leu122-Ser199-Arg426-Gly431		CAGETGAAGGAGAGCGTG CAGETGAAGGAGAGCGTG	
Lys121-Val200-Asn425-Lys432		CAGCTGAAGGAGAGCGTG	
Vall20-Ile201-Ile424-Ala433		CAGCTGAAGGAGAGCGTG	
110201-116121-N10133	(019)	ANNA VALUE AUGUST A LO	AUAUTHUUNE A

WO 00/39303	32	/ 65 PCT/US99/31272
Val12U-11e2U1B-11e424-Ala433	(619)	
Consensus		CAGCTGAAGGAGAGCGTGGAGATCAACTGC
	, ,	691 720
Leul22-Serl99-Tryp427-Gly431	(661)	ACCCGCCCAACAACAACACCCGCAAGAGC
Val127-Asn195-Arg426-Gly431	(691)	ACCCGCCCAACAACACACCCGCAAGAGC
Val120-Thr202-Ile424-Ala433	(649)	
Leu122-Ser199-Arg426-Lys432		ACCCGCCCAACAACACCCCGCAAGAGC
Leu122-Ser199-Arg426-Gly431	(661)	ACCCCCCAACAACAACACCCGCAAGAGC
Lys121-Va1200-Asn425-Lys432	(655)	ACCEGCCCAACAACAACACCCGCAAGAGC
Val120-Ile201-Ile424-Ala433		ACCCGCCCAACAACAACACCCGCAAGAGC
Vall20-Ile201B-Ile424-Ala433		ACCCGCCCAACAACAACACCCGCAAGAGC
Consensus	(691)	ACCCGCCCAACAACAACACCCGCAAGAGC
Inv. 122 Com 100 March 427 Ct 421	4602.	721 750
Leu122-Ser199-Tryp427-Gly431	(691)	ATCACCATEGGECCCGGCCGCCCTTCTAC
Val127-Asn195-Arg426-Gly431 Val120-Thr202-Ile424-Ala433	(/21)	ATCACCATCGGCCCCGGCCGCCCTTCTAC
Leul22-Ser199-Arg426-Lys432		ATCACCATCGGCCCGGCCGCCTTCTAC
Leu122-Ser199-Arg426-Gly431	(691)	ATCACCATCGGCCCCGGCGCCCTTGTAC ATCACCATCGGCCCCGCGCGCGCCTTGTAC
Lys121-Val200-Asn425-Lys432	(685)	ATCACCATCGGCCCGGCCGCCCTTCTAC
Val120-Ile201-Ile424-Ala433	(679)	ATCACCATCGGCCCGGCCGCCTTCTAC
Val120-Ile201B-Ile424-Ala433	(679)	ATCACCATCGGCCCGGCCGCCCTTCTAC
Consensus		ATCACCATCGGCCCGGCCGCCTTCTAC
	( /	751 780
Leu122-Ser199-Tryp427-Gly431	(721)	GCCACCGGCGACATCGCCGC
Val127-Asn195-Arg426-Gly431	(751)	
Val120-Thr202-Ile424-Ala433	(709)	GCCACCGGCGACATCACCGCCGACATCCGC
Leu122-Ser199-Arg426-Lys432	(721)	GCCACCGCGACATCATCGGCGACATCCGC
Leu122-Ser199-Arg426-Gly431	(721)	GCCACCGGCGACATCATCGGCGACATCCGC
Lys121-Va1200-Asn425 <b>-</b> Lys432	(715)	GCCACEGGCGACATCATEGGCGACATCEGC
Val120-Ile201-Ile424-Ala433	(709)	GCCACGGCGACATCATCGGCGACATCCGC
Val120-Ile201B-Ile424-Ala433	(709)	GCCACGGGGACATCATCGGCGACATCGGC
Consensus	(751)	GCCACCGGCGACATCATCGGCGACATCCGC
Leu122-Ser199-Tryp427-Gly431	(751)	781 810
Vall27-Asn195-Arg426-Gly431	(751)	CAGGGGGAGAAG
Val120-Thr202-Ile424-Ala433	(781) (739)	CAGGCCCACTGCAACATCAGCGGCGAGAAG
Leu122-Ser199-Arg426-Lys432	(751)	CAGGOCACTOCAACATCAGCGCGCGAGAAG CAGGOCACTGCAACATCAGCGGGGAGAAG
Leu122-Ser199-Arg426-Gly431	(751)	CACGOZOAC CCARCATCAGCGGGGACAGG
Lys121-Val200-Asn425-Lys432	(745)	CAGGECEACTGCACATCAGCGGCGAGAAG
Val120-Ile201-Ile424-Ala433	(739)	CAGGCCCACTGCAACATCAGCGGCGAGAAG
Val120-Ile201B-Ile424-Ala433	(739)	CAGGCCCACTGCAACATCAGCGGCGAGAAG
Consensus		CAGGCCCACTGCAACATCAGCGGCGAGAAG
	••	811 840
Leu122-Ser199-Tryp427-Gly431	(781)	TEGAACACACECTGAAGCAGATCGTGACC
Val127-Asn195-Arg426-Gly431	(811)	TGERACAACACCCTGAAGCAGATCGTGACC
Val120-Thr202-Ile424-Ala433	(769)	TGCAACACCCCTGAAGCAGATCGTGACC
Leu122-Ser199-Arg426-Lys432		TGGAACACACCCTGAAGCAGATCGTGACC
Leu122-Ser199-Arg426-Gly431		TGGAACACCCTGAAGCAGATCGTGACC
Lys121-Val200-Asn425-Lys432		TGGAACACCCTGAAGCAGATCGTGACC
Val120-Ile201-Ile424-Ala433		TGGAACAAGACCCTGAAGCAGATCGTGACC
Vall20-Ile201B-Ile424-Ala433		TGGAACAACACCCTGAAGCAGATCGTGACC
Consensus	(811)	TGGAACACCCTGAAGCAGATCGTGACC
Leu122-Cor100-m427 01423	(011)	841 870
Leu122-Ser199-Tryp427-Gly431 Val127-Asn195-Arg426-Gly431		AAGCTGCAGGCCCAGTTCGGCAACAAGACC
Vall20-Thr202-Ile424-Ala433		AAGCTGCAGGCCCAGTTCGGCAACAAGACC
Leu122-Ser199-Arg426-Lys432		ÄÄGCTGCÄGGCCCAGTTCGGCAACAÄGACC ÄÄGCTGCÄGGCCCAGTTCGGCAACAAGACC
octiss nig420-by3432	(011)	www.rannoncount.concunctificities

WO 00/39303		PCT/US99/31272
	33	./ <b>65</b>
Leu122-Ser199-Arg426-Gly431	(811)	AAGCTGCAGGCCCAGTTCGGCAACAAGACC
Lys121-Va1200-Asn425-Lys432 Va1120-Ile201-Ile424-Ala433	(805)	
Vall20-11e201-11e424-A[a433 Vall20-11e201B-I1e424-Ala433	(799)	
Consensus	(799)	
Consensus	(841)	
Leu122-Ser199-Tryp427-Gly431	(841)	871 900 ATCGTGTTCAAGCAGAGCAGCGGGGGGAC
Val127-Asn195-Arg426-Gly431	(871)	ATCGTGTTCAAGCAGCAGCGGCGCGAC
Val120-Thr202-Ile424-Ala433		ATCGTETTCAAGCAGAGCAGCGGCGAC
Leu122-Ser199-Arg426-Lys432	(841)	ATCGTGTTCAAGCAGAGCAGCGGCGGCGAC
Leu122-Ser199-Arg426-Gly431	(841)	ATCGTGTTCAAGCAGAGCAGCGGCGGCGAC
Lys121-Val200-Asn425-Lys432	(835)	ATCGTGTTCAAGCAGAGCAGCGGCGGCGAC
Val120-Ile201-Ile424-Ala433	(829)	ATCGTGTTCAAGCAGAGCAGCGGCGCGAC
Vall20-Ile201B-Ile424-Ala433	(829)	ATCGTGTTCAAGCAGAGCAGCGGCGGCGAC
Consensus	(871)	ATCGTGTTCAAGCAGAGCAGCGGCGGCGAC
		901 930
Leul22-Ser199-Tryp427-Gly431	(871)	CCCGAGATCGTGATGCACAGCTTCAACTGC
Val127-Asn195-Arg426-Gly431	(901)	CCCGAGATCGTGATGCACAGCTTCAACTGC
Val120-Thr202-Ile424-Ala433		CCCGAGATCGTGATGCACAGCTTCAACTGC
Leu122-Ser199-Arg426-Lys432	(871)	CCCGAGATCGTGATGCACAGCTTCAACTGC
Leu122-Ser199-Arg426-Gly431	(871)	CCCGAGATCGTGATGCACAGCTTCAACTGC
Lys121-Val200-Asn425-Lys432 Val120-Ile201-Ile424-Ala433	(865)	CCCGAGATCGTGATGCACAGCTTCAAGTGC
Val120-11e201-11e424-A1a433 Val120-11e201B-11e424-Ala433	(859) (859)	CCCGAGATCGTGATGCACAGCTTGAACTGC
Consensus		CCCGAGATCGTGATGCACAGCTTCAACTGC CCCGAGATCGTGATGCACAGCTTCAACTGC
Consensus	()01)	931 960
Leu122-Ser199-Tryp427-Gly431	(901)	GGCGGCGAGTYCTTCTACTGCAACAGCACC
Val127-Asn195-Arg426-Gly431	(931)	GGCGGGGACTTCTTCTACTGCAACAGCACC
Val120-Thr202-Ile424-Ala433	(889)	GGCGCGAGTICTTCTACTGCAACAGCACC
Leu122-Ser199-Arg426-Lys432	(901)	GGGGGGAGTICTTCTACTGCAACAGCACC
Leu122-Ser199-Arg426-Gly431	(901)	GGCGGGAGTTCTTCTACTGCAACAGGACC
Lys121-Val200-Asn425-Lys432	(895)	GGCGGCGAGTTCTTCTACTGCAACAGCACC
Val120-Ile201-Ile424-Ala433	(889)	GGCGGGGAGUICATCTACTGCAACAGCACC
Val120-Ile201B-Ile424-Ala433	(889)	CERCES CATERVIOUS CONTROL OF COALCACE
Consensus	(931)	GGCGGCGAGTTCTTCTACTGCAACAGCACC
T		961 990
Leu122-Ser199-Tryp427-Gly431	(931)	CAGCTGTTEPACAGGACCTGGAACAACACC
Val127-Asn195-Arg426-Gly431 Val120-Thr202-Ile424-Ala433	(961)	CAGCEGUTCAACAGGACCTGGAACACACC
Leu122-Ser199-Arg426-Lys432	(919)	CAGCTETTCAACAGCACCTGGAACAACACC
Leu122-Ser199-Arg426-Gly431		CAGCTGTTCAACAGGACCTGGAACACACC
Lys121-Val200-Asn425-Lys432	(925)	<u>FRECTETTCANTAGEACTTGGAACAACT</u> CC <u>CAGCTETTCANC</u> AGCACGTGGAACAACACC
Val120-Ile201-Ile424-Ala433		CAGCUGITEARCAGCACCTGGAACACCC
Val120-Ile201B-Ile424-Ala433		CACCUSTICAN CACCACCIGGARCACTIC
Consensus	(961)	CAGCTGTTCAACAGCACCTGGAACAACACC
		991 1020
Leul22-Ser199-Tryp427-Gly431	(961)	ATICGESSCEAASAACACCAACGGGACCATC
Val127-Asn195-Arg426-Gly431		ATCGGCCCCAACAACACCAACGGCACCATC
Val120-Thr202-Ile424-Ala433	(949)	ATCGGCCCGACACACCCACGGCACCATC
Leu122-Ser199-Arg426-Lys432	(961)	ATCGGCCCAACAACACCAACGGCACCATC
Leu122-Ser199-Arg426-Gly431	(961)	ATCGGCCCCAACAACACCAACGGCACCATC
Lys121-Val200-Asn425-Lys432		ATCGGCCCAACAACAACGGCACGATC
Val120-Ile201-Ile424-Ala433		ATCGGCCCAACAACACCAACGGCACCATC
Val120-Ile201B-Ile424-Ala433		ATCGGCCCAACAACACCAACGGCACCATC
Consensus		ATCGGCCCCAACACACCAACGGCACCATC
Leu122-Ser199-Tryp427-Gly431		1021 1050
Trypaza-Giyasi	())11	ACCCTGCCCTGCCGCATCAAGCAGATCATC

```
Val127-Asn195-Arg426-Gly431
                                (1021) ACCCTGCCCTGCCGCATCAAGCAGATCATC
 Val120-Thr202-Ile424-Ala433
                                (979) ACCCTGCCCTGCCGCATCAAGCAGATCATC
 Leu122-Ser199-Arg426-Lvs432
                                (991) ACCCTGCCCTGCCGCATCAAGCAGATCATC
 Leu122-Ser199-Arg426-Gly431
                                (991) ACCCTGCCTGCCGCATCAAGCAGATCATC
 Lys121-Val200-Asn425-Lys432
                                (985) ACCCTGCCCTGCCGCATCAGCAGATCATC
                                (979) ACCCTGCCGCATCAAGCAGATCATC
 Val120-Ile201-Ile424-Ala433
Val120-Ile201B-Ile424-Ala433
                                (979) ACCCTGCCCTGCCGCATCAGCAGATCATC
                    Consensus
                               (1021) ACCCTGCCCTGCCGCATCAAGCAGATCATC
                                      1051
Leu122-Ser199 Tryp427-Gly431
                               (1021) AACCGCTGGGGCGCAAGGCCATGTACGCC
 Val127-Asn195-Arg426-Gly431
                               (1051) AACCGCGGCGGCGGCAAGGCCATGTACGCC
                               (1009) ------GCGGG---GCCATGTACGCC
 Val120-Thr202-Ile424-Ala433
                               (1021) NACCGCGGCGGCAACAAGGCCATGTACGCC
 Leu122-Ser199-Arg426-Lys432
                               (1021) AACCGCGGCAGCGGCAAGGCCATGTACGCC
 Leu122-Ser199-Arg426-Gly431
                               (1015) AAC-----GCCCCCAAGGCCATGTAGGCC
 Lys121-Val200-Asn425-Lys432
 Val120-Ile201-Ile424-Ala433
                               (1009)
                                      -----GCCGGC---GCCATGYAGGCC
                               (1009) -----GCCGCC---GCCATGTACGCC
Vall20-Ile201B-Ile424-Ala433
                               (1051) AACCGC G GGCGGCAAGGCCATGTACGCC
                    Consensus
                                      1081
                               (1051) CCCCCCATCCGCGCCCAGATCCGCTGCAGC
Leu122-Ser199 Tryp427-Gly431
 Val127-Asn195-Arg426-Gly431
                               (1081) CCCCCCA COGCGGCCAGA CCGC GCAGC
                               (1027) CCCCCCATCCCCCGCCAGAVICCCCCCCAGC
 Val120-Thr202-Ile424-Ala433
                               (1051) CCCCCCATCCGCCCAGATCCGCTGCAGC
 Leu122-Ser199-Arg426-Lys432
                               (1051) CCCCCATCCGCGCCAGATCCGCTGCAGC
 Leu122-Ser199-Arg426-Gly431
                               (1039) CCCCCCATCGCGGCCACATCCGCTGCAGC (1027) CCCCCCATGCGGGGCCACATCCGGTGCAGC
 Lys121-Val200-Asn425-Lys432
 Val120-Ile201-Ile424-Ala433
Val120-Ile201B-Ile424-Ala433
                               (1027) CCCCCCATCCGCGGCCAGATCCGCTGCAGC
                               (1081) CCCCCATCCGCGGCCAGATCCGCTGCAGC
                                      1111
                                                                 1140
                               (1081) AGCAACATCACCGGCCTGCTGCTGACCCGC
Leu122-Ser199 Tryp427-Gly431
                               (1111) AGCARCATGACCGGCCTGCTGACCGGC
 Val127-Asn195-Arg426-Gly431
                               Val120-Thr202-Ile424-Ala433
 Leu122-Ser199-Arg426-Lys432
                               (1081) AGEALER LEAGUEGE GEORGE CONCERNE AGE GE
 Leu122-Ser199-Arg426-Gly431
 Lys121-Val200-Asn425-Lys432
                               (1069)
                                      AGGNACATE RESEGGETTECTS CTS ACCESC
 Val120-Ile201-Ile424-Ala433
                               (1057)
                                      ीटोकः ४१८ कः १४१७% (टाटाटाटाटाटा विकास सम्बद्धाः स्टार्थः टा<mark>ट</mark>
Val120-Ile201B-Ile424-Ala433
                               (1057) ACOMICA MICOGGO CHICONE SECOGO
                   Consensus
                               (1111) AGCAACATCACCGGCCTGCTGACCCGC
                                                                 1170
Leu122-Ser199 Tryp427-Gly431
                               (1111) GAUGGEGERAGGAGATCAGEAAGACCACC
                                      CANCECCEUNACEACATION CONTRACONOC
 Val127-Asn195-Arg426-Gly431
                               (1141)
                               (1087)
                                      daerdescribedadrescribede
 Val120-Thr202-Ile424-Ala433
                                      CALCOCCUBERGGAGATICAGCALICAGE
Leu122-Ser199-Arg426-Lys432
                               (1111)
                                      #)/##Cocses:(@#)((1/4/#)(Coss(w)(#)/#
Leu122-Ser199-Arg426-Gly431
                               (11111)
                               (1099)
 Lys121-Val200-Asn425-Lys432
                                      GACGGCGGGAAGGAGATCAGCAACACCAGC
 Val120-Ile201-Ile424-Ala433
                               (1087)
                                      GACGGCGGCAAGGAGATCAGCAAGACCAGC
                               (1087)
                                      GACGGCGGCAAGGAGATCAGCAACACCACC
Val120-Ile201B-Ile424-Ala433
                   Consensus
                                      GACGGCGCAAGGAGATCAGCAACACCACC
Leu122-Ser199 Tryp427-Gly431
                               (1141) GAGATETTECGCCCGGCGGCGGCGACATG
 Val127-Asn195-Arg426-Gly431
                               (1171)
                                      GAGATETTECGCCCCGGCGGCGCCGACATG
Val120-Thr202-Ile424-Ala433
                                      GAGATETTECGCCCCGGCGGCGGCGACATG
                               (1117)
Leu122-Ser199-Arg426-Lys432
                               (1141)
                                      GAGATETTECGCCCCGGCGCGCGCGACATG
                                      GAGATETTECCCCCCGCCGCGCGCGACATG
Leu122-Ser199-Arg426-Gly431
                               (1141)
Lys121-Val200-Asn425-Lys432
                               (1129)
                                      GAGATCTTCCGCCCCGGCGCGCGCCACATG
Val120-Ile201-Ile424-Ala433
                               (1117)
                                      GAGATETTCCGCCCCGGCGCGCGACATG
Val120-Ile201B-Ile424-Ala433
                               (1117) GAGATETTCCGCCCCGGCGCGCGCGACATG
```

Consensus	(1171)	GAGATCTTCCGCCCCGGCGCGCGCGACATG
		1201 1230
Leul22-Serl99 Tryp427-Gly431	(1171)	#decamoner replication content of
Val127-Asn195-Arg426-Gly431	(1201)	COCEAGA ACTES CECAGO LA COTETA CA A
Val120-Thr202-Ile424-Ala433	(1147)	COLUMN A STREET COLUMN COLUMN COTTOTA CANC
Leu122-Ser199-Arg426-Lys432	(1171)	SENSEASARATICS COAGCOCCACCACCACACACACA
Leu122-Ser199-Arg426-Gly431	(1171)	A STATION TO SECURE OF SECURE ACTIONS
Lys121-Va1200-Asn425-Lys432	(1159)	POCCACALCECTCCCCACCAGAGACCTCCTACAAC
Val120-Ile201-Ile424-Ala433	(1147)	CGCGACAACTGCCTGCACCGACCTCTACBAG
Val120-Ile201B-Ile424-Ala433	(1147)	COMMEANCING CONCRETE TOTAL AND
Consensus	(1201)	CGCGACAACTGGCGCGGGCGGGCTGTACAAG
		1231 1260
Leu122-Ser199 Tryp427-Gly431	(1201)	UALAAGGROGI GAAGATGGAGCCCC TCGGC
Val127-Asn195-Arg426-Gly431	(1231)	VACABESTICS SANGATICGACOCCUTS (IS-
Val120-Thr202-Ile424-Ala433	(1177)	TACABAGE GARGATO CACOTOCO
Leu122-Ser199-Arg426-Lys432	(1201)	TACAAGGTGGTGAAGATGGAECCCCTGGGG
Leu122-Ser199-Arg426-Gly431	(1201)	TACAAGGTEGTGAARATCGAGCCCCTGGGC
Lys121-Val200-Asn425-Lys432	(1189)	TACARGET CARGATCGAGCCCTGGGC
Val120-Ile201-Ile424-Ala433	(1177)	TACAMGGIOSTGANGATCGAGCCCCTGGGC
Vall20-Ile201B-Ile424-Ala433	(1177)	TACAAGGTGBEGAAGATCGAGCGCCTGGGC
Consensus	(1231)	TACAAGGTGGTGAAGATCGAGCCCCTGGGC
		1261 1290
Leul22-Ser199 Tryp427-Gly431	(1231)	GTGDCCCCCACCAAGCCCAAGCGCCCCCTG
Val127-Asn195-Arg426-Gly431	(1261)	GTGGCCCCCAGCAAGGCCZAGGGCCGCGTG
Val120-Thr202-Ile424-Ala433	(1207)	GIGGOCCCOACCNAGGCCAAGGGCGGCGCGTG
Leu122-Ser199-Arg426-Lys432	(1231)	GTGGCCCCCACCAAGGCCCAAGGCCCCCCCATG
Leu122-Ser199-Arg426-Gly431	(1231)	GTGGCCCCCAGEAAGGCCCAAGCCCCCGCGTG
Lys121-Val200-Asn425-Lys432	(1219)	GTGGCCCCACCAAGGCCCAAGGGCCCGCGTG
Vall20-Ile201-Ile424-Ala433	(1207)	GTGGCCCCACCAMGCCCAMGCGCCGCGTG
Vall20-Ile201B-Ile424-Ala433	(1207)	STGSECCCEACCAAGGCCAAGGGCGCGTG
Consensus	(1261)	GTGGCCCCACCAAGGCCAAGCGCCGCGTG
I a v 1 2 2 C a m 1 0 0 M 4 2 7 C 3 4 2 3	(1061)	1291 1320
Leu122-Ser199 Tryp427-Gly431	(1261)	GTGCACCGCGAGAAGCGCGCGTGACCCTG
Val127-Asn195-Arg426-Gly431	(1291)	GTGGAGCGCGAGAAGCGCGCGCGACCCTG
Val120-Thr202-Ile424-Ala433	(1237)	STGCBGCGCGAERAGCGCGCGCGCTGACCLTG
Leu122-Ser199-Arg426-Lys432	(1261)	STGCAGCGCLAS#AGEGEECGGTCACCLTG
Leul22-Ser199-Arg426-Gly431 Lys121-Val200-Asn425-Lys432	(1261)	GTGCAGCGCGAGAAGGGGGGCGTGACCCTG
Vall20-Ile201-Ile424-Ala433	(1249)	GTGCAGCGEGAGAGGGGGCGCGTCACCCTG
Val120-Ile2018-Ile424-Ala433	(1237)	GT GCAPAC ENGLAGGE AGGECT CELL CALCET G
	(1237)	GTGCAGCGGGGGAGAGGGTGACCGTG
Consensus	(1291)	GTGCAGCGCGAGAAGCGCGCCGTGACCCTG
Leu122-Ser199 Tryp427-Gly431	/12011	1321 - 1350
Vall27-Asn195-Arg426-Gly431	(1291) (1321)	GCGCATOTHECTCGCTTCCTGCCCCC
Vall20-Thr202-Ile424-Ala433	(1267)	GGC CARCANGETS GCCTTTCTGCCC SCC
Leu122-Ser199-Arg426-Lys432	(1291)	GCC BEAUTHE TOGGET TO STEEL TO C
Leu122-Ser199-Arg426-Gly431		GCCDP-ARCHIDECTEGGDDT(GD)ACGCGCC
Lys121-Val200-Asn425-Lys432	(1291) (1279)	GGCG-CATGTS CTGGGCTTCETGGGCCC
Val120-Ile201-Ile424-Ala433	(1267)	GG GCANGTICCIGGGTTTCCTG-GCGCC
Val120-Ile2018-Ile424-Ala433	(1267)	GG GG ATTE TE TE GG CTT COT G G G G C C GG G GC AT G FT C CT G G C T T C T G G C G C C
Consensus	(1321)	
Consensus	(1321)	GGCGCCATGTTCCTGGGCTTCCTGGGCGCC 1351 1380
Leu122-Ser199 Tryp427-Gly431	(1321)	1351 1380 GCGGGAGCAS PAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG
Val127-Asn195-Arg426-Gly431	(1351)	GCCGC/AGCACTA TGGGCCTCCCTACCCTG
Val120-Thr202-Ile424-Ala433	(1297)	GOOGLASSA CATEGORGO CA CAGOCAG
Leu122-Ser199-Arg426-Lys432	(1321)	GCCCPPACCAC PROCESCE CONTROL OF CONTROL
Leu122-Ser199-Arg426-Gly431	(1321)	SCCGGCACGACGATGGGGGCCCGCTGCCTG
	, /	

WO 00/39303	36 /	65 PCT/US99/31272
Lys121-Val200-Asn425-Lys432	(1309)	GCCGGCAGCACCATGGGCGCCCGCAGCCTG
Val120-Ile201-Ile424-Ala433	(1297)	
Val120-Ile201B-Ile424-Ala433	(1297)	
Consensus	(1351)	GCCGCAGCACCATGGGCGCCCGCAGCCTG
Consensus	(1331)	1381 1410
Leu122-Ser199 Tryp427-Gly431	(1351)	ACCOTGACCGTGCAGGCCCGCCAGCTGCTG
Val127-Asn195-Arg426-Gly431		
	(1381)	ACCCTGAÇCĞTGCAGGCCCGCCAGGTGCTG
Val120-Thr202-Ile424-Ala433		ACCCTGACCGTGCAGGCCCGCCAGCTGCTG
Leu122-Ser199-Arg426-Lys432	(1351)	ACCCTGACCGTGCAGGCCCGCCAGCTGCTG
Leu122-Ser199-Arg426-Gly431	(1351)	ACCCTGACCGTGCAGGCCCGCCAGCTGCTG
Lys121-Val200-Asn425-Lys432	(1339)	ACCCTGACCCTGCAGGCCCGCCAGCTGCTG
Vall20-Ile201-Ile424-Ala433	(1327)	ACCCTGACCGTGCAGCCCGCCAGCTGCTG
Val120-Ile201B-Ile424-Ala433	(1327)	ACCCTGACCGTGCAGGCCCGCCAGCTGCTG
Consensus	(1381)	ACCCTGACCGTGCAGGCCCGCCAGCTGCTG
		1411 1440
Leu122-Ser199 Tryp427-Gly431	(1381)	AGCGGCATCGTGCAGCAGCAGCAACCTG
Val127-Asn195-Arg426-Gly431	(1411)	AGCGGCATCGTGCAGCAGCAGCAGCCTG
Val120-Thr202-Ile424-Ala433	(1357)	AGCGGCATCGTGCAGCAGCAGCACCTG
Leu122-Ser199-Arg426-Lys432	(1381)	AGCGGCATCGTGCAGCAGCAGCAACCTG
Leu122-Ser199-Arg426-Gly431	(1381)	AGCGGCATCGTGCAGCAGCAGCAACCTG
Lys121-Va1200-Asn425-Lys432	(1369)	AGCGGCATCGTGCAGCAGCAGCACCACCTG
Val120-Ile201-Ile424-Ala433	(1357)	AGCGGCATCGTGCAGCAGCAGCACCACCTG
Val120-Ile201B-Ile424-Ala433	(1357)	AGCGGCATCGTGCAGCAGCAGCAACCACCTG
Consensus	(1411)	AGCGGCATCGTGCAGCAGCAGCAACCTG
		1441 1470
Leu122-Ser199 Tryp427-Gly431	(1411)	CTGCGCGCCATEGAGGCCCAGCAGCACCTG
Val127-Asn195-Arg426-Gly431	(1441)	CTGCGCGCCATCGAGGCCCAGCAGCACCTG
Val120-Thr202-Ile424-Ala433	(1387)	CTGCGCGCCATCGAGGCCCCAGCAGCACCTG
Leu122-Ser199-Arg426-Lys432	(1411)	CTGCGCGCATCGAGGCCCAGCAGCACCTG
Leu122-Ser199-Arg426-Gly431	(1411)	CTGCGCGCATCGAGGCCCAGCAGCACCTG
Lys121-Val200-Asn425-Lys432	(1399)	CTGCGCGCCATCGAGGCCCCAGCAGCACCTG
Val120-Ile201-Ile424-Ala433	(1387)	CTGCGCGCATGGAGGCCCCAGCAGCACCTG
Val120-Ile201B-Ile424-Ala433	(1387)	CTGGGGGCATCGAGGCCCAGCAGCACCTG
Consensus	(1441)	CTGCGCGCCATCGAGGCCCAGCAGCACCTG
		1471 1500
Leu122-Ser199 Tryp427-Gly431	(1441)	EXECUTE VE VE CHEMICE CHEMICA A COME
Val127-Asn195-Arg426-Gly431	(1471)	enternyeleur in schleggeren en sternyteerte
Val120-Thr202-Ile424-Ala433		OTGCACOTOAC COTGTGGGGGAVCAAGCAG
Leu122-Ser199-Arg426-Lys432	(1441)	PINGENCETONER PROPRIES CONTRACTOR
Leu122-Ser199-Arg426-Gly431	(1441)	CHGGAGGREACGERGTGGGGGGATGAAGGAG
Lys121-Val200-Asn425-Lys432	(1429)	CHGGAGERGAC & GHGRGEGGGAYCAAGGAG
Val120-Ile201-Ile424-Ala433	(1417)	MICO/CONCANDACHONGEGGOANIAAAGOAG
Val120-Ile201B-Ile424-Ala433		CTGCAGCTGACCGTGTGGGGCATCAAGCAG
Consensus		CTGCAGCTGACCGTGTGGGGCATCAAGCAG
Conscisus	(13/1)	1501 1530
Leu122-Ser199 Tryp427-Gly431	(1471)	कार-अः(सटस्थरटस्टर्स्ट्रस्ट्रस्ट्रस्ट्रस्ट्रस्ट
Val127-Asn195-Arg426-Gly431		CTGCAGGCCCCGGGTGCTGGGGGGG
Val120-Thr202-Ile424-Ala433	(1301)	GHG #AGGGGAGGAGGAGGG
Leu122-Ser199-Arg426-Lys432		Diseases secure construction
Leu122-Ser199-Arg426-Gly431 Lys121-Va1200-Asn425-Lys432		ETCAGGE SEGRECITEGE GTGGAGGE C
	(1439)	CTGCAGGGCGCGTGCTGGAGGGC
Val120-Ile201-Ile424-Ala433		CTGCAGGCCGCGTGGTGGAGGGC
Vall20-Ile201B-Ile424-Ala433		CTGCAGGGCGGGGTGCTGGCCGTGGAGCGC
Consensus	(1501)	CTGCAGGCCGCGTGCTGGAGCGC
7100 o 100 m (0m m)	(1501)	1531 1560
Leu122-Ser199 Tryp427-Gly431		TACCTGAAGGACCAGCAGCTGCTGGGCATC
Val127-Asn195-Arg426-Gly431	(1531)	TACCTGAAGGACCAGCAGCTGCTGGGCATC

WO 00/39303	38 /	65	PCT/US99/31272
		1711	1740
Leu122-Ser199 Tryp427-Gly431	(1681)	ATCGAGGAGAGCCA	GAACCAGCAGAGAAG
Val127-Asn195-Arg426-GIy431	(1711)	ATCGAGGAGAGCCA	GAACCAGCAGGAGAAG
Val120-Thr202-Ile424-Ala433	(1657)	ATCGAGGAGAGCCA	GAACCAGCAGGAGAAG
Leu122-Ser199-Arg426-Lys432	(1681)		GAACCAGCAGGAGAAG
Leu122-Ser199-Arg426-Gly431	(1681)		GAACCAGCAGGAGAAG
Lys121-Val200-Asn425-Lys432	(1669)		GAACCAGCAGGAGAAG
Val120-Ile201-Ile424-Ala433	(1657)		GAACCAGCAG <b>GAGA</b> AG
Vall20-Ile201B-Ile424-Ala433	(1657)		GAACCAGCAGGAGAAG
Consensus	(1711)		GAACCAGCAGGAGAAG 1770
1122 Cam100 Mm427 Cl421	(1711)	1741	GCTGGAGCTGGACAAG
Leu122-Ser199 Tryp427-Gly431	(1711)		dia and a second and a second and a second a sec
Val127-Asn195-Arg426-Gly431 Val120-Thr202-Ile424-Ala433	(1741) (1687)		GCTGGAGCTGGACAAG GCTGGAGCTGGACAAG
Leu122-Ser199-Arg426-Lys432	(1711)		GCTGGAGCTGGACAAG
Leu122-Ser199-Arg426-Gly431	(1711)		CCTGGAGCTGGACAAG
Lys121-Val200-Asn425-Lys432	(1699)		GCTGGAGCTGGA <b>CA</b> AG
Val120-Ile201-Ile424-Ala433	(1687)	AACGAGCAGGAGCT	ectggagctggag <b>a</b> ag
Val120-Ile201B-Ile424-Ala433	(1687)		GCTGGAGCTGGACAAG
Consensus	(1741)	And the second control of the section of the second control of the	GCTGGAGCTGGACAAG
	•	1771	1800
Leu122-Ser199 Tryp427-Gly431	(1741)	TGEGCCAGCCTGTG	GAACTGGTTCGACATC
Val127-Asn195-Arg426-Gly431	(1771)	ricecce checericite	GAACTGGTTCGACATC
Val120-Thr202-Ile424-Ala433	(1717)		gaactggttcgacatc
Leu122-Ser199-Arg426-Lys432	(1741)		gaactggtte <b>gaca</b> tc
Leu122-Ser199-Arg426-Gly431	(1741)		GAACTGGTTCGACATC
Lys121-Val200-Asn425-Lys432	(1729)		GAACTGGTT <b>CGALA</b> TC
Val120-Ile201-Ile424-Ala433	(1717)		GAACTGGTTCGACATC
Vall20-Ile201B-Ile424-Ala433	(1717)	AND SECTION AND PROPERTY OF THE PARTY.	GAACTGGTTCGACATC
Consensus	(1771)		GAACTGGTTCGACATC
Tau 122 Car 100 May 427 Clu421	(1771)	1801	1830
Leu122-Ser199 Tryp427-Gly431	(1771)		GTACATCAAGATCTTC
Val127-Asn195-Arg426-Gly431 Val120-Thr202-Ile424-Ala433	(1801) (1747)		GIVACATICAAGATEETIC SIVACATICAAGATEETIC
Leu122-Ser199-Arg426-Lys432	(1771)		GTACATCAAGATOTTC
Leu122-Ser199-Arg426-Gly431	(1771)		CIVA CANTO ARC ASSOCIATE
Lys121-Val200-Asn425-Lys432	(1759)		CTACATORACATORIC
Val120-Ile201-Ile424-Ala433	(1747)		CITA CANCAL GANG UIC
Val120-Ile201B-Ile424-Ala433	(1747)		GTACATCAAGATCTTC
Consensus	(1801)		GTACATCAAGATCTTC
		1831	1860
Leu122-Ser199 Tryp427-Gly431	(1801)	AVEAUGANCE/IGGG	ecceptice tecches
Val127-Asn195-Arg426-Gly431	(1831)	ATOMICAN CITEGO	CGCCTGGTGGGGCTG
Val120-Thr202-Ile424-Ala433	(1777)		GGCCTGGTGGGCGTG
Leu122-Ser199-Arg426-Lys432	(1801)		GGCCTGSTGGG CAG
Leu122-Ser199-Arg426-Gly431	(1801)		CGCCTGGTGGGCCTG
Lys121-Va1200-Asn425-Lys432			GGGETGGTGGGGGG
Val120-Ile201-Ile424-Ala433			eggeetgeteggetig
Vall20-Ile201B-Ile424-Ala433	(1777)		eggergergegerg
Consensus	(1831)		CGGCCTGGTGGGCCTG
Tau 122 Ga-100 M : 1407 G1: 427	(1071)	1861	1890
Leu122-Ser199 Tryp427-Gly431	(1831)		CGTGCTGAGCATCGTG
Val127-Asn195-Arg426-Gly431	(1861)		CGTGCTGAGCATCGTG
Val120-Thr202-Ile424-Ala433	(1807) (1831)		CGTGCTGAGCATEGTG
Leu122-Ser199-Arg426-Lys432 Leu122-Ser199-Arg426-Gly431	(1831)		CGTGCTGAGCATCGTG CGTGCTGAGCATCGTG
Lys121-Val200-Asn425-Lys432	(1819)		CGTGCTGAGCATCGTG
-1-1-1 -GTEOO-USHIED DASIDE	, /		

```
Val120-Ile201-Ile424-Ala433
                                      (1807) CGCATCGTGTTCACCGTGCTGAGCATCGTG
  Val120-Ile201B-Ile424-Ala433
                                      (1807) CGCATCGTGTTCACCGTGCTGAGCATCGTG
                        Consensus
                                      (1861) CGCATCGTGTTCACCGTGCTGAGCATCGTG
 Leu122-Ser199 Tryp427-Gly431
                                      (1861) AACCGCGTGCGCCAGGGCTACAGCCCCCCTG
  Val127-Asn195-Arg426-Gly431
                                     (1891) AACCGCGTGCGCCAGGGCTACAGCCCCCTG
  Val120-Thr202-Ile424-Ala433
                                     (1837) AACCGCGTGCGCCAGGGCTACAGCCCCCTG
  Leu122-Ser199-Arg426-Lys432
                                     (1861) AACCGCGTGCGCCAGGGCTACAGCCCCCTG
(1861) AACGGCGTGCGCCCAGGGCTACAGCCCCGTG
  Leu122-Ser199-Arg426-Glv431
  Lys121-Val200-Asn425-Lys432
                                     (1849) AACCGCGTGCGCCAGGGCTACAGCCCCCTG
  Val120-Ile201-Ile424-Ala433
                                     (1837) AACCGCGTGCGCCAGGGCTACAGCCCCCTG
 Val120-Ile201B-Ile424-Ala433
                                     (1837) AACCGCGTGCGCCAGGGCTACAGCCCCCTG
                                     (1891) AACCGCGTGCGCCAGGGCTACAGCCCCCTG
                        Consensus
 Leu122-Ser199 Tryp427-Gly431
                                     (1891) AGCTTCCAGACCCGCTTCCCCGCCCCCGGC
  Val127-Asn195-Arg426-Gly431
                                     (1921) AGCTTCCAGACCCGGTTCCCGGCCCCCGC
                                     (1867) AGCTTCCAGACCCGCTTCCCCGGCCCGCGC
  Val120-Thr202-Ile424-Ala433
  Leu122-Ser199-Arg426-Lys432
                                     (1891) AGCTTCCAGACEGGCTTCCCGGCCCCGCCC
  Leu122-Ser199-Arg426-Gly431
                                     Lys121-Val200-Asn425-Lys432
                                     (1879) AGCITCCAGACCEGCTTCCCCGCCCCCCC
  Val120-Ile201-Ile424-Ala433
                                     (1867) AGCTICCAGACCCGCTTCCGCGCCCCCGC
 Val120-Ile201B-Ile424-Ala433
                                     (1921) AGCTTCCAGACCCGCTTCCCCGCCCCCGC
                       Consensus
 Leu122-Ser199 Tryp427-Gly431
                                     (1921) GGCCCGACGCCCGAGGGCATCGAGGAG
  Val127-Asn195-Arg426-Gly431
                                     Val120-Thr202-Ile424-Ala433
                                             GGCCCGACCGCCCGGACGCCATGCACGAG
                                     (1897)
 Leu122-Ser199-Arg426-Lys432
                                     (1921)
                                            GGCCCCGACCGCCCCGACGGCATCCACCAC
 Leu122-Ser199-Arg426-Gly431
                                            GGeteGGNess assert/title+telect/lactic
                                     (1921)
 Lys121-Val200-Asn425-Lys432
                                     (1909)
                                             द्वित्वसम्बद्धाः द्वार्ट्स्बब्द्धस्य । द्वारा प्रदेशस्य ।
 Val120-Ile201-Ile424-Ala433
                                    (1897)
                                             विवर्तक एक १० वर्ष तर वर्षकार सिन्त मार्ग स्टेस्टर है।
Val120-Ile201B-Ile424-Ala433
                                    (1897)
                                            ਜ਼ਫ਼ਫ਼ਫ਼ਜ਼ਖ਼ਜ਼੶ਫ਼ਜ਼ਫ਼ਫ਼ਖ਼ਫ਼ਫ਼ਜ਼ੑ੶ਫ਼ਜ਼ਜ਼ਜ਼੶ਖ਼ਫ਼ਜ਼੶੶ਫ਼ਜ਼੶ਫ਼
                       Consensus
                                    (1951) GGCCCCGACCGCCCGAGGGCATCGAGGAG
Leu122-Ser199 Tryp427-Gly431
                                            स्थायम्बद्धसम्बद्धसम्बद्धसम्बद्धसम्बद्धसम्बद्धसम्बद्धसम्बद्धसम्बद्धसम्बद्धसम्बद्धसम्बद्धसम्बद्धसम्बद्धसम्बद्धस
                                    (1951)
 Val127-Asn195-Arg426-Gly431
                                    (1981)
                                            GAGGGGGGAGAGAGGGAGGGGAGGG
 Val120-Thr202-Ile424-Ala433
                                    (1927)
                                            त्तर्भवत्वत्वत्वयम् एवद्दद्वतम् ४००० वर्गम् ४५८० ह
 Leu122-Ser199-Arg426-Lys432
                                    (1951)
                                            द्वभूद्रदर्भ सद्भवस्था अवदर्शन स्थापन स्
 Leu122-Ser199-Arg426-Gly431
                                            (1951)
 Lys121-Val200-Asn425-Lys432
                                            वर्गस्यव्हार्यसम्बद्धाः स्थापन्य स्थापन्य स्थापन्य ।
वर्गस्यक्षार्थसम्बद्धाः स्थापन्य स्थापन्य स्थापन्य स्थापन्य स्थापन्य स्थापन्य स्थापन्य स्थापन्य स्थापन्य स्थापन
                                    (1939)
 Val120-Ile201-Ile424-Ala433
                                    (1927)
                                            त्रसम्बर्भन्यः । स्वत्रस्य । स्वत्रस्य
Val120-Ile201B-Ile424-Ala433
                                            तः,(तर्म्यम्नित्रसः,(प्रधानित्सः,(र्वभावनः,(ब्रह्मस्
                                    (1927)
                       Consensus
                                    (1981)
                                            GAGGGCGAGCGCGACCGCAGC
Leu122-Ser199 Tryp427-Gly431
                                    (1981)
                                            ्रित्रव्यव्यवक्रम्यात्मम् । अस्य स्वित्यम् । अस्य स्वित्यम् । अस्य स्वित्यम् । अस्य स्वित्यम् । अस्य स्वित्यम्
 Val127-Asn195-Arg426-Glv431
                                    (2011)
                                            Victor Andrical Cassectoring Chickers of
 Val120-Thr202-Ile424-Ala433
                                    (1957)
                                            AGCCOCKECKECACEGGERICHER
 Leu122-Ser199-Arg426-Lys432
                                    (1981) AGCCCCTGGTGCACGGCCTGCTGCGCCCTG
 Leu122-Ser199-Arg426-Gly431
                                    (1981) AGCCCCTGGTGGACGGCGTGGTGGCCCTG
 Lys121-Val200-Asn425-Lys432
                                    (1969) AGCOSCENES VGC ACGGG VGC AGGOCOTG
(1957) AGCCCCCTGG GCACGGGC GTGC AGCTG
 Val120-Ile201-Ile424-Ala433
Val120-Ile201B-Ile424-Ala433
                                    (1957) AGCCCCTGGTGCACGGCCTGGTGGCCCTG
                                    (2011) AGCCCCTGGTGCACGGCCTGCTGGCCCTG
                      Consensus
                                            2041
Leu122-Ser199 Tryp427-Gly431
                                   (2011) Andrese Andrese Section (2011) Andrese Grand Section (2011) Andrese Grand Section (2011)
 Val127-Asn195-Arg426-Gly431
 Val120-Thr202-Ile424-Ala433
```

2221

2250

WO 00/39303	41	/ 65	PCT/US99/31272
Leu122-Ser199 Tryp427-Gly431	(2191)	GACGCCAN	en e avi e e e con e e e e e e e e e e e e e e e e e e e
Val127-Asn195-Arg426-Glv431	(2221)	GACGCCAT	CGCEAFEGEGGTGGCCGAGGGC
Val120-Thr202-Ile424-Ala433	(2167)	GACGCCAT	SEGUATUCCOGTEGECGAGGCC
Leu122-Ser199-Arg426-Lys432	(2191)		ece vegees cook
Leu122-Ser199-Arg426-Gly431	(2191)	GACGCCAT	seenviregeecheeeeveec
Lys121-Val200-Asn425-Lys432	(2179)	GACGCCAT	SGCPATEGECGTGGCGGC
Val120-Ile201-Ile424-Ala433	(2167)	GACGCCAT	SECCATEGOCGTGGCCCAGGGC
Val120-Ile201B-Ile424-Ala433	(2167)	GACGCCAT	CGCCATCGCCGTGGCCGAGGGC
Consensus	(2221)	GACGCCAT	CGCCATCGCCGTGGCCGAGGGC
		2251	2280
Leu122-Ser199 Tryp427-Gly431	(2221)	ACCGACCG	PATCATCGAGGTGGCCCAGCGC
Val127-Asn195-Arg426-Gly431	(2251)	ACCGACCG	TATCATCGAGGTGGCCCAGCGC
Val120-Thr202-Ile424-Ala433	(2197)	ACGGACCG	ATCATCGAGGTGGCCCAGGGC
Leu122-Ser199-Arg426-Lys432	(2221)	ACEGAGCG	ATCATOGAGGTGGCCCAGCGC
Leu122-Ser199-Arg426-Gly431	(2221)	ACCGACCG	CATCATEGAGGTGGCCCAGCGC
Lys121-Val200-Asn425-Lys432	(2209)	ACCGACCG	ATCATEGAGGTGGCCCAGEGC
Val120-Ile201-Ile424-Ala433	(2197)		ATCATCGAGGTGGCCCAGGGC
Val120-Ile201B-Ile424-Ala433	(2197)	ACCGACCG	ATCATEGAGGTGGCCCAGCGC
Consensus	(2251)		CATCATCGAGGTGGCCCAGCGC
Tou122-For100 Mm-427 Cl-421	(0051)	2281	2310
Leu122-Ser199 Tryp427-Gly431	(2251)	ARCEGECE	GENTAL TECACATION CETEC
Val127-Asn195-Arg426-Gly431 Val120-Thr202-Ile424-Ala433	(2281)	Avicelege	GECTTECTGCACATECCCCGC
Leu122-Ser199-Arg426-Lys432	(2227)	ALLSGERGE	CCCTTCCTGCACATCCCCGCC
Leu122-Ser199-Arg426-Eys432	(2251) (2251)	ATCGGCCGG	<u>Belligetigeacatereces</u> e
Lys121-Val200-Asn425-Lys432	(2231)	WI CONTROLLED	GEETTECTGCACATECCCCCC
Val120-Ile201-Ile424-Ala433	(2227)	ALL DOLLOS	GCTTCCTGCACATCGCCCGC
Val120-Ile201B-Ile424-Ala433	(2227)	NTO CACACO	GCCTTCCTGCACATCCCCCGC GCCTTCCTGCACATCCCCCGC
Consensus	(2281)		GCCTTCCTGCACATCCCCGC
0050545	(2201)	2311	2340
Leu122-Ser199 Tryp427-Gly431	(2281)		PACCCP WCCACCCCCCCC
Val127-Asn195-Arg426-Gly431	(2311)	CCCATCCC	engedenteensecseessys
Val120-Thr202-Ile424-Ala433	(2257)	CCO TON	GARCO BEACHARD COCKETAS
Leu122-Ser199-Arg426-Lys432	(2281)	MC100 11 50 00 00	ल <i>्ट्रस्ट स्टब्स्ट सम्मान</i> सम्बद्धाः स्टब्स्ट स्टब्स्ट स्टब्स्ट स्टब्स्ट स्टब्स्ट स्टब्स्ट स्टब्स्ट स्टब्स्ट स
Leu122-Ser199-Arg426-Gly431	(2281)	6 E 6 . 1 10 0 0 0 0	<b>७</b> ः(सराम्बर्ध <b>र्यः (स्ट्</b> रावस <b>्</b> र्वे स्ट्रास्
Lys121-Val200-Asn425-Lys432	(2269)		CARGE CON GAGE GOOGLE
Val120-Ile201-Ile424-Ala433	(2257)		er/teletesminer/teletist news (c
Val120-Ile201B-Ile424-Ala433	(2257)		ଷ୍ଟ୍ରପ୍ରକ୍ରେଷ୍ଟ୍ରକ୍ୟକ୍ରକ୍ରକ୍ତ୍ର G
Consensus	(2311)		CAGGGCTTCGAGCGCGCCCTG
			352
Leu122-Ser199 Tryp427-Gly431	(2311)	shift) 6.4.16830	GAG
Val127-Asn195-Arg426-Gly431	(2341)	entervirent	CAG
Val120-Thr202-Ile424-Ala433	(2287)	chile, s. c. (mile	
Leu122-Ser199-Arg426-Lys432	(2311)	Sancon artonac	GAN'S
Leu122-Ser199-Arg426-Gly431	(2311)	estackilistically	970
Lys121-Val200-Asn425-Lys432		CECTAGE	
Val120-Ile201-Ile424-Ala433	(2287)	CANCEL PROPERTY AND ADMIC	
Val120-Ile201B-Ile424-Ala433	(2287)	CIGIAACIC	SAC
Consensus	(2341)	CTGTAACTC	GAĞ

#### **SEQ ID NO:3 VAL120-ALA204**

GTCTTCGTTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTG TGGAAGGAGGCCACCACCACCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT GCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCCAGGAGATCGTGCT GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGGGCGCCGGCGCCTGCCCCAA GGTGAGCTTCGAGCCCATCCCATCCACTACTGCGCCCCGCGGCTTCGCCATCCTGAAGTG CAACGACAAGAAGTTCAACGGCAGCGGCCCCTGCACCAACGTGAGCACCGTGCAGTGCACCC ACGGCATCCGCCCGTGGTGAGCACCCAGCTGCTGCTGAACGGCAGCCTGGCCGAGGAGGGC GTGGTGATCCGCAGCGAGAACTTCACCGACAACGCCAAGACCATCATCGTGCAGCTGAAGGA GAGCGTGGAGATCAACTGCACCCGCCCCAACAACAACACCCGCAAGAGCATCACCATCGGCC CCGGCCGCGCCTTCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCCACTGCAACA TCAGCGGCGAGAAGTGGAACACACCCTGAAGCAGATCGTGACCAAGCTGCAGGCCCAGTTC GGCAACAGACCATCGTGTTCAAGCAGAGCAGCGGCGACCCCGAGATCGTGATGCACAG CTTCAACTGCGGCGGGGGTTCTTCTACTGCAACAGCACCCAGCTGTTCAACAGCACCTGGAA CAACACCATCGGCCCCAACACACCCAACGGCACCATCACCCTGCCCTGCCGCATCAAGCAGA TCATCAACCGCTGGCAGGAGGTGGGCAAGGCCATGTACGCCCCCCCATCCGCGGCCAGATC CGCTGCAGCAGCAACATCACCGGCCTGCTGCTGACCCGCGACGGCGAAGGAGATCAGCAA CACCACCGAGATCTTCCGCCCGGCGGCGGCGACATGCGCGACAACTGGCGCAGCGAGCTGT ACAAGTACAAGGTGGTGAAGATCGAGCCCCTGGGCGTGGCCCCACCAAGGCCAAGCGCCGC GTGGTGCAGCGCGAGAAGCGCGCCGTGACCCTGGGCGCCATGTTCCTGGGCTTCCTGGGCGCC GCCGCAGCACCATGGGCGCCCGCAGCCTGACCCTGACGTGCAGGCCCGCCAGCTGCTGAG CGGCATCGTGCAGCAGCAGCACCACCTGCTGCGCGCCCATCGAGGCCCAGCAGCACCTGCTGC AGCTGACCGTGTGGGGCATCAAGCAGCTGCAGGCCCGCGTGCTGGCCGTGGAGCGCTACCTG AAGGACCAGCAGCTGCTGGGCATCTGGGGCTGCAGCGCCAAGCTGATCTGCACCACCGCCGT GCCTGGAACGCCAGCTGGAGCAACAAGAGCCTGGACCAGATCTGGAACAACATGACCTGGA TGGAGTGGGAGCGCGAGATCGACAACTACACCAACCTGATCTACACCCTGATCGAGGAGAGC CAGAACCAGCAGGAGAAGAACGAGCAGGAGCTGCTGGAGCTGGACAAGTGGGCCAGCCTGT GGAACTGGTTCGACATCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCATGATCGTGGGCG GCCTGGTGGGCCTGCGCATCGTGTTCACCGTGCTGAGCATCGTGAACCGCGTGCGCCAGGGCT ACAGCCCCTGAGCTTCCAGACCCGCTTCCCCGCCCCCGGGCCCCGACCGCCCCGAGGGCA TCGAGGAGGAGGCGCGAGCGCGACCGCACCGCAGCCCCCTGGTGCACGGCCTGCTG GCCTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTCAGCTACCACCGCCTGCGCGACCTG ATCCTGATCGCCGCCGCATCGTGGAGCTGCTGGGCCGCCGCGGCTGGGAGGCCCTGAAGTAC TGGGGCAACCTGCTGCAGTACTGGATCCAGGAGCTGAAGAACAGCGCCGTGAGCCTGTTCGA CGCCATCGCCATCGCCGTGGCCGAGGGCACCGACCGCATCATCGAGGTGGCCCAGCGCATCG GCCGCGCTTCCTGCACATCCCCGCCGCATCCGCCAGGGCTTCGAGCGCCCCTGCTGTAAC **TCGAG** 

#### SEQ ID NO:4 VAL120-ILE201

GTCTTCGTTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTG TGGAAGGAGGCCACCACCACCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT GCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCCAGGAGATCGTGCT GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGGGCGGCATCACCCAGGCCTG CCCCAAGGTGAGCTTCGAGCCCATCCCCTACTGCGCCCCGCCGGCTTCGCCATCCT GAAGTGCAACGACAAGAAGTTCAACGGCAGCGCCCCTGCACCAACGTGAGCACCGTGCAGT GCACCCACGGCATCCGCCCGTGGTGAGCACCCAGCTGCTGCAACGGCAGCCTGGCCGAG GAGGGCGTGGTGATCCGCAGCGAGAACTTCACCGACAACGCCAAGACCATCATCGTGCAGCT GAAGGAGAGCGTGGAGATCAACTGCACCCGCCCCAACAACACCCCGCAAGAGCATCACCA TCGGCCCGGCCGCCTTCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCCACT GCAACATCAGCGGCGAGAAGTGGAACAACACCCTGAAGCAGATCGTGACCAAGCTGCAGGCC CAGTTCGGCAACAAGACCATCGTGTTCAAGCAGAGCAGCGGCGGCGACCCCGAGATCGTGAT GCACAGCTTCAACTGCGGCGGCGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAACAGCAC CTGGAACAACACCATCGGCCCCAACAACACCAACGGCACCATCACCCTGCCCTGCCGCATCA CAGATCCGCTGCAGCAGCAACATCACCGGCCTGCTGCTGACCCGCGACGGCGACGGCAAGGAGAT CAGCAACACCACCGAGATCTTCCGCCCCGGCGGCGGCGACATGCGCGACAACTGGCGCAGCG AGCTGTACAAGTACAAGGTGGTGAAGATCGAGCCCCTGGGCGTGGCCCCACCAAGGCCAAG CGCCGCGTGTGCAGCGCGAGAAGCGCGCCGTGACCCTGGGCGCATGTTCCTGGGCTTCCTG GGCGCCGCCGGCAGCACCATGGGCGCCCGCAGCCTGACCCTGACCGTGCAGGCCCGCCAGCT GCTGAGCGCATCGTGCAGCAGCAGAACAACCTGCTGCGCCCATCGAGGCCCAGCAGCACC TGCTGCAGCTGACCGTGTGGGGCATCAAGCAGCTGCAGGCCCGCGTGCTGGCCGTGGAGCGC TACCTGAAGGACCAGCAGCTGCTGGGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCAC CGCCGTGCCCTGGAACGCCAGCTGGAGCAACAAGAGCCTGGACCAGATCTGGAACAACATGA CCTGGATGGAGTGGGAGCGCGAGATCGACAACTACACCAACCTGATCTACACCCTGATCGAG GAGAGCCAGAACCAGCAGGAGAAGAACGAGCAGGAGCTGGAGCTGGACAAGTGGGCCA GCCTGTGGAACTGGTTCGACATCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCATGATCG TGGGCGGCCTGGTGGCCTGCGCATCGTGTTCACCGTGCTGAGCATCGTGAACCGCGTGCGCC AGGGCTACAGCCCCTGAGCTTCCAGACCCGCTTCCCCGCCCCCGCGGCCCCGACCGCCCCG AGGGCATCGAGGAGGAGGGCGGCGAGCGCGACCGCAGCAGCCCCCTGGTGCACGG CCTGCTGGCCCTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTCAGCTACCACCGCCTGCG CGACCTGATCCTGATCGCCGCCCGCATCGTGGAGCTGCTGGGCCGCCGCGGCTGGGAGGCCCT GAAGTACTGGGGCAACCTGCTGCAGTACTGGATCCAGGAGCTGAAGAACAGCGCCGTGAGCC TGTTCGACGCCATCGCCATCGCCGTGGCCGAGGGCACCGCATCATCGAGGTGGCCCAGC GCATCGCCGCCCTTCCTGCACATCCCCCGCCACATCCGCCAGGGCTTCGAGCGCGCCCTGC TGTAACTCGAG

## SEQ ID NO:5 VAL120-ILE201B

GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGTGTGGAGCAGTCTTCG TTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTGTGGAAGGAGGCCA CCACCACCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGTGCACAACGTGTGGGCCACCC ACGCCTGCGTGCCCACCGACCCCAACCCCCAGGAGATCGTGCTGGAGAACGTGACCGAGAACTTCAACA TGTGGAAGAACACATGGTGGAGCAGATGCACGAGGACATCATCAGCCTGTGGGACCAGAGCCTGAAGC CCTGCGTGCCCGGCATCACCCAGGCCTGCCCCAAGGTGAGCTTCGAGCCCATCCCCATCCACTACTGCGC CCCCGCCGGCTTCGCCATCCTGAAGTGCAACGACAAGAAGTTCAACGGCAGCGGCCCCTGCACCAACGT GAGCACCGTGCAGTGCACCCACGGCATCCGCCCCGTGGTGAGCACCCAGCTGCTGCTGAACGGCAGCCT GGCCGAGGAGGGCGTGGTGATCCGCAGCGAGAACTTCACCGACAACGCCAAGACCATCATCGTGCAGCT GAAGGAGAGCGTGGAGATCAACTGCACCCGCCCCAACAACAACACCCGCAAGAGCATCACCATCGGCCC CGGCCGCGCTTCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCCACTGCAACATCAGCGGC GAGAAGTGGAACAACACCCTGAAGCAGATCGTGACCAAGCTGCAGGCCCAGTTCGGCAACAAGACCATC GTGTTCAAGCAGAGCAGCGGCGGCGACCCCGAGATCGTGATGCACAGCTTCAACTGCGGCGGCGAGTTC TTCTACTGCAACAGCACCCAGCTGTTCAACAGCACCTGGAACAACACCATCGGCCCCAACAACACCAAC GGCACCATCACCCTGCCCTGCCGCATCAAGCAGATCATCAACCGCTGGCAGGAGGTGGGCAAGGCCATG TACGCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGCAACATCACCGGCCTGCTGCTGACCCGCGACG GCGGCAAGGAGATCAGCAACACCACCGAGATCTTCCGCCCCGGCGGCGCGACATGCGCGACAACTGGC GCAGCGAGCTGTACAAGTACAAGGTGGTGAAGATCGAGCCCCTGGGCGTGGCCCCCACCAAGGCCAAGC GCCGCGTGGTGCAGCGCGAGAAGCGCGCCGTGACCCTGGGCGCCATGTTCCTGGGCTTCCTGGGCGCCGC CGGCAGCACCATGGGCGCCCGCAGCCTGACCCTGACCGTGCAGGCCCGCCAGCTGCTGAGCGGCATCGT GCAGCAGCAGAACAACCTGCTGCGCGCCATCGAGGCCCAGCAGCACCTGCTGCAGCTGACCGTGTGGGG CATCAAGCAGCTGCAGGCCCGCGTGCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTGGGCAT CTGGGGCTGCAGCGCAAGCTGATCTGCACCACCGCCGTGCCCTGGAACGCCAGCTGGAGCAACAAGAG CCTGGACCAGATCTGGAACAACATGACCTGGATGGAGTGGGAGCGCGAGATCGACAACTACACCAACCT GATCTACACCCTGATCGAGGAGAGCCAGAACCAGCAGGAGAAGAACGAGCAGGAGCTGCTGGAGCTGG ACAAGTGGGCCAGCCTGTGGAACTGGTTCGACATCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCAT GATCGTGGGCGGCCTGGGGCCTGCGCATCGTGTTCACCGTGCTGAGCATCGTGAACCGCGTGCGCCAG GGCTACAGCCCCTGAGCTTCCAGACCCGCTTCCCCGCCCCCCGCGGCCCCGACCGCCCCGAGGGCATCG AGGAGGAGGGCGCGACCGCGACCGCAGCAGCCCCCTGGTGCACGGCCTGCTGGCCCTGATCT GGGACGACCTGCGCAGCCTGTGCCTGTTCAGCTACCACCGCCTGCGCGACCTGATCCTGATCGCCGCCCG CATCGTGGAGCTGCTGGGCCGCCGCGGCTGGGAGGCCCTGAAGTACTGGGGCAACCTGCTGCAGTACTG GATCCAGGAGCTGAAGAACAGCGCCGTGAGCCTGTTCGACGCCATCGCCATCGCCGTGGCCGAGGGCAC CGACCGCATCATCGAGGTGGCCCAGCGCATCGGCCGCCCTTCCTGCACATCCCCCGCCGCATCCGCCAG GGCTTCGAGCGCGCCCTGCTGTAACTCGAGCGTGCT

FIG. 8

### **SEQ ID NO:6 LYS121-VAL200**

GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGTGTGGAGCA GTCTTCGTTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTG TGGAAGGAGGCCACCACCACCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT GCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCCCAGGAGATCGTGCT GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGAAGGCCCCCGTGATCACCCA GGCTGCCCAAGGTGAGCTTCGAGCCATCCCCATCCACTACTGCGCCCCGCCGGCTTCGC CATCCTGAAGTGCAACGACAAGAAGTTCAACGGCAGCGGCCCCTGCACCAACGTGAGCACCG TGCAGTGCACCCACGGCATCCGCCCCGTGGTGAGCACCCAGCTGCTGCTGAACGGCAGCCTGG CCGAGGAGGCGTGGTGATCCGCAGCGAGAACTTCACCGACAACGCCAAGACCATCATCGTG CAGCTGAAGGAGAGCGTGGAGATCAACTGCACCCGCCCAACAACAACACCCGCAAGAGCAT CACCATCGGCCCGGCCGCCTTCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGC CCACTGCAACATCAGCGGCGAGAAGTGGAACAACACCCTGAAGCAGATCGTGACCAAGCTGC AGGCCCAGTTCGGCAACAAGACCATCGTGTTCAAGCAGAGCAGCGGCGGCGACCCCGAGATC GTGATGCACAGCTTCAACTGCGGCGGCGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAAC AGCACCTGGAACAACACCATCGGCCCCAACAACACCCAACGGCACCATCACCCTGCCCTGCCG CATCAAGCAGATCATCAACCGCTGGCAGGAGGTGGGCAAGGCCATGTACGCCCCCCCATCC GCGGCCAGATCCGCTGCAGCAGCAACATCACCGGCCTGCTGCTGACCCGCGACGGCGGCAAG GAGATCAGCAACACCACCGAGATCTTCCGCCCCGGCGGCGGCGACACTGCGCGACAACTGGCG CAGCGAGCTGTACAAGTACAAGGTGGTGAAGATCGAGCCCCTGGGCCGTGGCCCCCACCAAGG CCAAGCGCCGCGTGCCGCGCGAGAAGCGCGCCGTGACCCTGGGCGCCATGTTCCTGGGC TTCCTGGGCGCCGCCGGCAGCACCATGGGCGCCCGCAGCCTGACCCTGACCGTGCAGGCCCGC CAGCTGCTGAGCGGCATCGTGCAGCAGCAGAACAACCTGCTGCGCGCCATCGAGGCCCAGCA GCACCTGCTGCAGCTGACCGTGTGGGGCATCAAGCAGCTGCAGGCCCGCGTGCTGGCCGTGG AGCGCTACCTGAAGGACCAGCAGCTGCTGGGCATCTGGGGCTGCAGCGGCAAGCTGATCTGC ACCACCGCCGTGCCCTGGAACGCCAGCTGGAGCAACAAGAGCCTGGACCAGATCTGGAACAA CATGACCTGGATGGAGTGGGAGCCCGAGATCGACAACTACACCAACCTGATCTACACCCTGA TCGAGGAGAGCCAGAACCAGCAGGAGAAGAACGAGCAGGAGCTGCTGGAGCTGGACAAGTG GGCCAGCCTGTGGAACTGGTTCGACATCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCAT GATCGTGGGCGGCCTGGGCCTGCGCATCGTGTTCACCGTGCTGAGCATCGTGAACCGCGT GCGCCAGGGCTACAGCCCCTGAGCTTCCAGACCGCTTCCCCGCCCCCCGCGGCCCGACCG CCCCGAGGGCATCGAGGAGGAGGGCGGCGAGCGCGACCGCAGCAGCAGCCCCTGGTGC ACGGCTGCTGGCCTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTCAGCTACCACCGCC TGCGCGACCTGATCCTGATCGCCGCCCGCATCGTGGAGCTGCTGGGCCGCCGCGGCTGGGAGG CCCTGAAGTACTGGGGCAACCTGCTGCAGTACTGGATCCAGGAGCTGAAGAACAGCGCCGTG CAGCGCATCGGCCGCCTTCCTGCACATCCCCCGCCGCATCCGCCAGGGCTTCGAGCGCGCC CTGCTGTAACTCGAGCGTGCT

FIG. 9

**SEQ ID NO:7: LEU122-SER199** 

GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGTGTGGAGCA GTCTTCGTTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTG TGGAAGGAGGCCACCACCACCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT GCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCCAGGAGATCGTGCT GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGAAGCTGGGCAACAGCGTGAT CACCCAGGCCTGCCCCAAGGTGAGCTTCGAGCCCATCCCCATCCACTACTGCGCCCCGCCGG CTTCGCCATCCTGAAGTGCAACGACAAGAAGTTCAACGGCAGCGGCCCCTGCACCAACGTGA GCACCGTGCAGTGCACCCACGGCATCCGCCCCGTGGTGAGCACCCAGCTGCTGCTGAACGGC AGCCTGGCCGAGGGGGGCGTGGTGATCCGCAGCGAGACTTCACCGACAACGCCAAGACCAT CATCGTGCAGCTGAAGGAGAGCGTGGAGATCAACTGCACCCGCCCAACAACAACACCCGCA AGAGCATCACCATCGGCCCGGCGCCTTCTACGCCACCGGCGACATCATCGGCGACATCC GCCAGGCCCACTGCAACATCAGCGGCGAGAAGTGGAACAACACCCTGAAGCAGATCGTGACC AAGCTGCAGGCCCAGTTCGGCAACAAGACCATCGTGTTCAAGCAGAGCAGCGGCGGCGACCC CGAGATCGTGATGCACAGCTTCAACTGCGGCGGCGAGTTCTTCTACTGCAACAGCACCCAGCT GTTCAACAGCACCTGGAACAACACCATCGGCCCCAACAACACCAACGGCACCATCACCTGC CCTGCCGCATCAAGCAGATCATCAACCGCTGGCAGGAGGTGGGCAAGGCCATGTACGCCCCC CCCATCCGCGGCCAGATCCGCTGCAGCAGCAACATCACCGGCCTGCTGCTGACCGGCGACGGC GGCAAGGAGATCAGCAACACCACCGAGATCTTCCGCCCCGGCGGCGGCGACATGCGCGACAA CTGGCGCAGCGAGCTGTACAAGTACAAGGTGGTGAAGATCGAGCCCCTGGGCGTGGCCCCCA CCAAGGCCAAGCGCGCGTGGTGCAGCGCGAGAAGCGCGCGTGACCCTGGGCGCCATGTTC CTGGGCTTCCTGGGCGCCGCCGCAGCACCATGGGCGCCCGCAGCCTGACCCTGACCGTGCAG GCCCGCCAGCTGCTGAGCGCATCGTGCAGCAGCAGCACCACCTGCTGCGCGCCATCGAGGC CCAGCAGCACCTGCTGCAGCTGACCGTGTGGGGCATCAAGCAGCTGCAGGCCCGCGTGCTGG CCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTGGGCATCTGGGGCTGCAGCGGCAAGCTG ATCTGCACCACCGCCGTGCCCTGGAACGCCAGCTGGAGCAACAAGAGCCTGGACCAGATCTG GAACAACATGACCTGGATGGAGTGGGAGCGCGAGATCGACAACTACACCAACCTGATCTACA CCCTGATCGAGGAGAGCCAGAACCAGCAGGAGAAGAACGAGCAGGAGCTGCTGGAGCTGGA CAAGTGGGCCAGCCTGTGGAACTGGTTCGACATCAGCAAGTGGCTGTGGTACATCAAGATCTT CATCATGATCGTGGGCCTGGTGGGCCTGCGCATCGTGTTCACCGTGCTGAGCATCGTGAA CCGCGTGCGCCAGGCCTACAGCCCCCTGAGCTTCCAGACCCGCTTCCCCGCCCCCCGCGGCCC CTGGTGCACGGCCTGCTGGCCCTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTCAGCTAC CACCGCCTGCGCGACCTGATCCTGATCGCCGCCCGCATCGTGGAGCTGCTGGGCCGCCGCGGC TGGGAGGCCCTGAAGTACTGGGGCAACCTGCTGCAGTACTGGATCCAGGAGCTGAAGAACAG GGTGGCCCAGCGCATCGGCCGCGCCTTCCTGCACATCCCCCGCCGCATCCGCCAGGGCTTCGA GCGCGCCTGCTGTAACTCGAGCGTGCT

FIG. 10

## SEQ ID NO:8 VAL120-THR202

GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGTGTGGAGCA GTCTTCGTTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTG TGGAAGGAGGCCACCACCACCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT GCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCCCAGGAGATCGTGCT GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG CCCCAAGGTGAGCTTCGAGCCCATCCCATCCACTACTGCGCCCCGCCGGCTTCGCCATCCT GAAGTGCAACGACAAGAAGTTCAACGGCAGCGCCCCTGCACCAACGTGAGCACCGTGCAGT GCACCCACGGCATCCGCCCCGTGGTGAGCACCCAGCTGCTGCAACGGCAGCCTGGCCGAG GAGGGCGTGGTGATCCGCAGCGAGAACTTCACCGACAACGCCAAGACCATCATCGTGCAGCT GAAGGAGAGCGTGGAGATCAACTGCACCCGCCCCAACAACACCCCGCAAGAGCATCACCA TCGGCCCGGCCGCCTTCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCCACT GCAACATCAGCGGCGAGAAGTGGAACAACACCCTGAAGCAGATCGTGACCAAGCTGCAGGCC CAGTTCGGCAACAAGACCATCGTGTTCAAGCAGAGCAGCGGCGGCGACCCCGAGATCGTGAT GCACAGCTTCAACTGCGGCGGCGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAACAGCAC CTGGAACACACCATCGGCCCCAACAACACCCAACGGCACCATCACCCTGCCCTGCCGCATCA AGCAGATCATCAACCGCTGGCAGGAGGTGGGCAAGGCCATGTACGCCCCCCCATCCGCGGC CAGATCCGCTGCAGCAGCAACATCACCGGCCTGCTGCTGACCCGCGACGGCGGCAAGGAGAT CAGCAACACCACCGAGATCTTCCGCCCCGGCGGCGGCGACATGCGCGACAACTGGCGCAGCG AGCTGTACAAGTACAAGGTGGTGAAGATCGAGCCCCTGGGCGTGGCCCCCACCAAGGCCAAG CGCCGCGTGGTGCAGCGCGAGAAGCGCGCGCGTGACCCTGGGCGCCATGTTCCTGGGCTTCCTG GGCGCCGCCGGCACCATGGGCGCCCGCAGCCTGACCCTGACCGTGCAGGCCCGCCAGCT GCTGAGCGGCATCGTGCAGCAGCAGAACAACCTGCTGCGCGCCATCGAGGCCCAGCAGCACC TGCTGCAGCTGACCGTGTGGGGCATCAAGCAGCTGCAGGCCCGCGTGCTGGCCGTGGAGCGC TACCTGAAGGACCAGCAGCTGCTGGGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCAC CGCCGTGCCCTGGAACGCCAGCTGGAGCAACAAGAGCCTGGACCAGATCTGGAACAACATGA CCTGGATGGAGTGGGAGCGCGAGATCGACAACTACACCAACCTGATCTACACCCTGATCGAG GAGAGCCAGAACCAGCAGGAGAAGAACGAGCAGGAGCTGCTGGAGCAAGTGGGCCA GCCTGTGGAACTGGTTCGACATCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCATGATCG TGGGCGGCCTGGTGGGCCTGCGCATCGTGTTCACCGTGCTGAGCATCGTGAACCGCGTGCGCC AGGGCTACAGCCCCTGAGCTTCCAGACCCGCTTCCCCGCCCCCGCGGCCCCGACCGCCCCG CCTGCTGGCCCTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTCAGCTACCACCGCCTGCG CGACCTGATCCTGATCGCCGCCCGCATCGTGGAGCTGCTGGGCCGCCGCGGGTGGGAGGCCCT GAAGTACTGGGGCAACCTGCTGCAGTACTGGATCCAGGAGCTGAAGAACAGCGCCGTGAGCC TGTTCGACGCCATCGCCGTGGCCGAGGGCACCGACCGCATCATCGAGGTGGCCCAGC GCATCGGCCGCCTTCCTGCACATCCCCCGCCGCATCCGCCAGGGCTTCGAGCGCGCCCTGC **TGTAACTCGAG** 

## **SEQ ID NO:9 TRP427-GLY431**

GTCTTCGTTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTG TGGAAGGAGGCCACCACCACCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT GCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCCAGGAGATCGTGCT GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGAAGCTGACCCCCTGTGCGTG ACCCTGCACTGCACCAACCTGAAGAACGCCACCAACACCCAAGAGCAGCAACTGGAAGGAGAT GGACCGCGGCGAGATCAAGAACTGCAGCTTCAAGGTGACCACCAGCATCCGCAACAAGATGC AGAAGGAGTACGCCCTGTTCTACAAGCTGGACGTGGTGCCCATCGACAACAACAACACCAGC TACAAGCTGATCAACTGCAACACCAGCGTGATCACCCAGGCCTGCCCCAAGGTGAGCTTCGA GCCCATCCCATCACTACTGCGCCCCCGCCGGCTTCGCCATCCTGAAGTGCAACGACAAGAA GTTCAACGGCAGCGCCCCTGCACCAACGTGAGCACCGTGCAGTGCACCCACGGCATCCGCC CCGTGGTGAGCACCCAGCTGCTGCAACGGCAGCCTGGCCGAGGAGGGCGTGGTGATCCGC AGCGAGAACTTCACCGACAACGCCAAGACCATCATCGTGCAGCTGAAGGAGAGCGTGGAGAT CAACTGCACCCGCCCAACAACACCCCGCAAGAGCATCACCATCGGCCCCGGCCGCCCT TCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCCACTGCAACATCAGCGGCGAG AAGTGGAACAACACCCTGAAGCAGATCGTGACCAAGCTGCAGGCCCAGTTCGGCAACAAGAC CATCGTGTTCAAGCAGAGCAGCGGCGGCGACCCCGAGATCGTGATGCACAGCTTCAACTGCG GCGGCGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAACAGCACCTGGAACAACACCATCG GCCCCAACACCCAACGGCACCATCACCCTGCCCTGCCGCATCAAGCAGATCATCAACCGCT GGGGCGCAAGGCCATGTACGCCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGCAACATC ACCGGCTGCTGCTGACCCGCGACGGCGCAAGGAGATCAGCAACACCACCGAGATCTTCCG CCCCGGCGGCGACATGCGCGACAACTGGCGCAGCGAGCTGTACAAGTACAAGTGGTGA AGATCGAGCCCTGGGCGTGGCCCCCACCAAGGCCAAGCGCCGCGTGGTGCAGCGCGAGAAG CGCGCCGTGACCCTGGGCGCATGTTCCTGGGCTTCCTGGGCGCCGCCGCAGCACCATGGGC GCCGCAGCCTGACCCTGACCGTGCAGGCCCGCCAGCTGCTGAGCGGCATCGTGCAGCAGCA GAACAACCTGCTGCGCGCCATCGAGGCCCAGCAGCACCTGCTGCAGCTGACCGTGTGGGGCA TCAAGCAGCTGCAGGCCCGCGTGCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTG GGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCGCCGTGCCCTGGAACGCCAGCTG ATCGACAACTACACCAACCTGATCTACACCCTGATCGAGGAGGCCAGAACCAGCAGGAGAA GAACGAGCAGGAGCTGCAGCTGGACCAGCCAGCCTGTGGAACTGGTTCGACATCA GCAAGTGGCTGTGGTACATCAAGATCTTCATCATGATCGTGGGCGGCCTGGTGGGCCTGCGCA TCGTGTTCACCGTGCTGAGCATCGTGAACCGCGTGCGCCAGGGCTACAGCCCCCTGAGCTTCC AGACCCGCTTCCCCGCCCCCCGGGCCCCGACCCCCCGAGGGCATCGAGGAGGAGGGCGGC GAGCGCGACCGCAGCAGCCCCCTGGTGCACGCCTGCTGGCCCTGATCTGGGACGA CCTGCGCAGCCTGTGCCTGTTCAGCTACCACCGCCTGCGCGACCTGATCCTGATCGCCGCCCG CATCGTGGAGCTGCTGGGCCGCCGCGGCTGGGAGCCCTGAAGTACTGGGGCAACCTGCTGC AGTACTGGATCCAGGAGCTGAAGAACAGCGCCGTGAGCCTGTTCGACGCCATCGCCATCGCC GTGGCCGAGGGCACCGACCGCATCATCGAGGTGGCCCAGCGCATCGGCCGCCCTTCCTGCA CATCCCCGCCGCATCCGCCAGGGCTTCGAGCGCCCCTGCTGTAACTCGAG

## **SEQ ID NO:10 ARG426-GLY431**

GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGTGTGGAGCA GTCTTCGTTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTG TGGAAGGAGGCCACCACCCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT GCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCCCAGGAGATCGTGCT GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGAAGCTGACCCCCCTGTGCGTG ACCCTGCACTGCACCTGAAGAACGCCACCAACACCAAGAGCAGCAACTGGAAGGAGAT GGACCGCGGCGAGATCAAGAACTGCAGCTTCAAGGTGACCACCAGCATCCGCAACAAGATGC AGAAGGAGTACGCCCTGTTCTACAAGCTGGACGTGGTGCCCATCGACAACGACAACACCAGC TACAAGCTGATCAACTGCAACACCAGCGTGATCACCCAGGCCTGCCCCAAGGTGAGCTTCGA GCCCATCCCATCCACTACTGCGCCCCGCCGGCTTCGCCATCCTGAAGTGCAACGACAAGAA GTTCAACGGCAGCGCCCCTGCACCAACGTGAGCACCGTGCAGTGCACCCACGGCATCCGCC CCGTGGTGAGCACCCAGCTGCTGAACGGCAGCCTGGCCGAGGAGGGCGTGGTGATCCGC AGCGAGAACTTCACCGACAACGCCAAGACCATCATCGTGCAGCTGAAGGAGAGCGTGGAGAT CAACTGCACCCGCCCAACAACAACACCCGCAAGAGCATCACCATCGGCCCCGGCCGCCCT TCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCCACTGCAACATCAGCGGCGAG AAGTGGAACAACACCCTGAAGCAGATCGTGACCAAGCTGCAGGCCCAGTTCGGCAACAAGAC CATCGTGTTCAAGCAGAGCAGCGGCGGCGACCCCGAGATCGTGATGCACAGCTTCAACTGCG GCGGCGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAACAGCACCTGGAACAACACCATCG GGCGGCGCAAGGCCATGTACGCCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGCAACAT CACCGGCCTGCTGACCCGCGACGGCGGCAAGGAGATCAGCAACACCACCGAGATCTTCC GCCCCGGCGGCGCGACATGCGCGACAACTGGCGCAGCGAGCTGTACAAGTACAAGGTGGTG AAGATCGAGCCCCTGGGCGTGGCCCCCACCAAGGCCAAGCGCCGCGTGGTGCAGCGCGAGAA GCGCGCCGTGACCCTGGGCGCCATGTTCCTGGGCTTCCTGGGCGCCGCCGGCAGCACCATGGG CGCCCGCAGCCTGACCCTGACCGTGCAGGCCCGCCAGCTGCTGAGCGGCATCGTGCAGCAGC AGAACAACCTGCTGCGCGCCATCGAGGCCCAGCAGCACCTGCTGCAGCTGACCGTGTGGGGC ATCAAGCAGCTGCAGGCCCGCGTGCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCT GGGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCGCCGTGCCCTGGAACGCCAGCT GATCGACAACTACACCAACCTGATCTACACCCTGATCGAGGAGGCCAGAACCAGCAGGAGA AGAACGAGCAGGAGCTGCAGACCAGCTGGGCCAGCCTGTGGAACTGGTTCGACATC ATCGTGTTCACCGTGCTGAGCATCGTGAACCGCGTGCGCCAGGGCTACAGCCCCCTGAGCTTC  ${\tt CGAGCGCGACCGCAGCAGCCCCCTGGTGCACGGCCTGGTCTGGCCCTGATCTGGGACG}$ ACCTGCGCAGCCTGTTCCAGCTACCACCGCCTGCGCGACCTGATCCTGATCGCCGCCC GCATCGTGGAGCTGCTGGGCCGCCGCGGCTGGGAGGCCCTGAAGTACTGGGGCAACCTGCTG CAGTACTGGATCCAGGAGCTGAAGAACAGCGCCGTGAGCCTGTTCGACGCCATCGCCATCGC CGTGGCCGAGGGCACCGACCGCATCATCGAGGTGGCCCAGCGCATCGGCCGCCTTCCTGC ACATCCCCGCCGCATCCGCCAGGGCTTCGAGCGCGCCCTGCTGTAACTCGAG

## **SEQ ID NO:11 ARG426-GLY431B**

GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGTGTGGAGCA GTCTTCGTTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTG TGGAAGGAGGCCACCACCACCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT GCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCCAGGAGATCGTGCT GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGAAGCTGACCCCCTGTGCGTG ACCCTGCACTGCACCAACCTGAAGAACGCCACCAACACCAAGAGCAGCAACTGGAAGGAGAT GGACCGCGGCGAGATCAAGAACTGCAGCTTCAAGGTGACCACCAGCATCCGCAACAAGATGC AGAAGGAGTACGCCCTGTTCTACAAGCTGGACGTGGTGCCCATCGACAACGACAACACCAGC TACAAGCTGATCAACTGCAACACCAGCGTGATCACCCAGGCCTGCCCCAAGGTGAGCTTCGA GCCCATCCCATCACTACTGCGCCCCGCCGGCTTCGCCATCCTGAAGTGCAACGACAAGAA GTTCAACGCAGCGCCCCTGCACCAACGTGAGCACCGTGCAGTGCACCCACGGCATCCGCC  ${\tt CCGTGGTGAGCACCCAGCTGCTGAACGGCAGCCTGGCCGAGGAGGGCGTGGTGATCCGC}$ AGCGAGAACTTCACCGACAACGCCAAGACCATCATCGTGCAGCTGAAGGAGAGCGTGGAGAT CAACTGCACCGCCCAACAACAACACCCGCAAGAGCATCACCATCGGCCCCGGCCGCCCT TCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCCACTGCAACATCAGCGGCGAG AAGTGGAACACCCTGAAGCAGATCGTGACCAAGCTGCAGGCCCAGTTCGGCAACAAGAC CATCGTGTTCAAGCAGAGCAGCGGCGGCGACCCCGAGATCGTGATGCACAGCTTCAACTGCG GCGCGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAACAGCACCTGGAACAACACCATCG GGCAGCGGCAAGGCCATGTACGCCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGCAACAT CACCGGCCTGCTGACCCGCGACGGCGCAAGGAGATCAGCAACACCACCGAGATCTTCC GCCCGGCGGCGACATGCGCGACAACTGCGCGCGCGAGCTGTACAAGTACAAGTTGGTG AAGATCGAGCCCCTGGGCGTGGCCCCCACCAAGGCCAAGCGCCGCGTGGTGCAGCGCGAGAA GCGCGCGTGACCCTGGGCGCCATGTTCCTGGGCTTCCTGGGCGCCGCCGGCAGCACCATGGG CGCCCGCAGCCTGACCCTGACCGTGCAGGCCCGCCAGCTGCTGAGCGGCATCGTGCAGCAGC AGAACAACCTGCTGCGCGCCATCGAGGCCCAGCAGCACCTGCTGCAGCTGACCGTGTGGGGC ATCAAGCAGCTGCAGGCCCGCGTGCTGGCCGTGCAGCTGCTGAAGGACCAGCAGCTGCT GGGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCGCCGTGCCCTGGAACGCCAGCT GATCGACAACTACACCAACCTGATCTACACCCTGATCGAGGAGAGCCAGAACCAGCAGGAGA AGAACGAGCAGGAGCTGGAGCTGGACAAGTGGGCCAGCCTGTGGAACTGGTTCGACATC AGCAAGTGGCTGTGGTACATCAAGATCTTCATCATGATCGTGGGCGGCCTGGTGGGCCTGCGC ATCGTGTTCACCGTGCTGAGCATCGTGAACCGCGTGCGCCAGGGCTACAGCCCCCTGAGCTTC CAGACCCGCTTCCCCGCCCCCCGGGCCCCGACGCCCCGAGGGCATCGAGGAGGAGGGGGG CGAGCGCGACCGCAGCAGCCCCCTGGTGCACGGCCTGCTGGCCCTGATCTGGGACG ACCTGCGCAGCCTGTTCCAGCTACCACCGCCTGCGCGACCTGATCCTGATCGCCGCCC GCATCGTGGAGCTGCTGGGCCGCCGCGGCTGGGAGGCCCTGAAGTACTGGGGCAACCTGCTG CAGTACTGGATCCAGGAGCTGAAGAACAGCGCCGTGAGCCTGTTCGACGCCATCGCCATCGC CGTGGCCGAGGGCACCGCATCATCGAGGTGGCCCAGCGCATCGGCCGCGCTTCCTGC ACATCCCCGCCGCATCCGCCAGGGCTTCGAGCGCGCCCTGCTGTAACTCGAG

#### **SEQ ID NO:12 ARG426-LYS432**

GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGTGTGGAGCA GTCTTCGTTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTG TGGAAGGAGGCCACCACCACCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT GCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCCAGGAGATCGTGCT GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGAAGCTGACCCCCCTGTGCGTG ACCCTGCACTGCACCAACCTGAAGAACGCCACCAACACCAAGAGCAGCAACTGGAAGGAGAT GGACCGCGGCGAGATCAAGAACTGCAGCTTCAAGGTGACCACCAGCATCCGCAACAAGATGC AGAAGGAGTACGCCCTGTTCTACAAGCTGGACGTGGTGCCCATCGACAACGACAACACCAGC TACAAGCTGATCAACTGCAACACCAGCGTGATCACCCAGGCCTGCCCCAAGGTGAGCTTCGA GCCCATCCCATCCACTACTGCGCCCCGCCGGCTTCGCCATCCTGAAGTGCAACGACAAGAA GTTCAACGGCAGCGCCCCTGCACCAACGTGAGCACCGTGCAGTGCACCCACGGCATCCGCC CCGTGGTGAGCACCCAGCTGCTGAACGCAGCCTGGCCGAGGAGGGCGTGGTGATCCGC AGCGAGAACTTCACCGACAACGCCAAGACCATCATCGTGCAGCTGAAGGAGAGCGTGGAGAT CAACTGCACCGCCCCAACAACACCCGCAAGAGCATCACCATCGGCCCGGCCGCGCCT TCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCCACTGCAACATCAGCGGCGAG AAGTGGAACAACACCCTGAAGCAGATCGTGACCAAGCTGCAGGCCCAGTTCGGCAACAAGAC CATCGTGTTCAAGCAGAGCAGCGGCGGCGACCCCGAGATCGTGATGCACAGCTTCAACTGCG GCGGCGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAACAGCACCTGGAACAACACCATCG GGCGGCAACAAGGCCATGTACGCCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGCAACAT CACCGCCTGCTGCTGACCCGCGACGCCGAAGGAGATCAGCAACACCACCGAGATCTTCC GCCCCGGCGGCGCGACATGCGCGACAACTGGCGCAGCGAGCTGTACAAGTACAAGGTGGTG AAGATCGAGCCCTGGGCGTGGCCCCACCAAGGCCAAGCGCCGTGGTGCAGCGCGAGAA GCGCGCGTGACCTGGGCGCCATGTTCCTGGGCTTCCTGGGCGCGCCGGCAGCACCATGGG CGCCCGCAGCCTGACCCTGACCGTGCAGGCCCGCCAGCTGCTGAGCGGCATCGTGCAGCAGC AGAACAACCTGCTGCGCGCCATCGAGGCCCAGCAGCACCTGCTGCAGCTGACCGTGTGGGGC ATCAAGCAGCTGCAGGCCCGCGTGCTGGCCGTGGAGCCTACCTGAAGGACCAGCAGCTGCT GGGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCGCCGTGCCCTGGAACGCCAGCT GATCGACAACTACACCAACCTGATCTACACCCTGATCGAGGAGGCCAGAACCAGCAGGAGA AGAACGAGCAGGAGCTGCAGCTGGACAAGTGGGCCAGCCTGTGGAACTGGTTCGACATC AGCAAGTGGCTGTGGTACATCAAGATCTTCATCATGATCGTGGGCGGCCTGGTGGGCCTGCGC ATCGTGTTCACCGTGCTGAGCATCGTGAACCGCGTGCGCCAGGGCTACAGCCCCCTGAGCTTC CAGACCCGCTTCCCCGCCCCCGGGCCCCGACGCCCCGAGGGCATCGAGGAGGAGGGCGG CGAGCGCGACCGCAGCAGCCCCCTGGTGCACGGCCTGCTGGCCCTGATCTGGGACG ACCTGCGCAGCCTGTGCCTGTTCAGCTACCACCGCCTGCGCGACCTGATCCTGATCGCCGCCC GCATCGTGGAGCTGCTGGGCCGCCGCGGCTGGGAGGCCCTGAAGTACTGGGGCAACCTGCTG CAGTACTGGATCCAGGAGCTGAAGAACAGCGCCGTGAGCCTGTTCGACGCCATCGCCATCGC CGTGGCCGAGGGCACCGCATCATCGAGGTGGCCCAGCGCATCGGCCGCGCCTTCCTGC ACATCCCCGCCGCATCCGCCAGGGCTTCGAGCGCGCCCTGCTGTAACTCGAG

#### SEQ ID NO:13 ASN425-LYS432

GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGTGTGGAGCA GTCTTCGTTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTG TGGAAGGAGGCCACCACCACCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT GCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCCCAGGAGATCGTGCT GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGAAGCTGACCCCCCTGTGCGTG ACCCTGCACTGCACCAACCTGAAGAACGCCACCAACACCAAGAGCAGCAACTGGAAGGAGAT GGACCGCGGCGAGATCAAGAACTGCAGCTTCAAGGTGACCACCAGCATCCGCAACAAGATGC AGAAGGAGTACGCCCTGTTCTACAAGCTGGACGTGGTGCCCATCGACAACGACAACACCAGC TACAAGCTGATCAACTGCAACACCAGCGTGATCACCCAGGCCTGCCCCAAGGTGAGCTTCGA GCCCATCCCATCCACTACTGCGCCCCGCCGCTTCGCCATCCTGAAGTGCAACGACAAGAA GTTCAACGCCAGCGCCCCTGCACCAACGTGAGCACCGTGCAGTGCACCCACGGCATCCGCC CCGTGGTGAGCACCCAGCTGCTGCAACGGCAGCCTGGCCGAGGAGGGCGTGGTGATCCGC AGCGAGAACTTCACCGACAACGCCAAGACCATCATCGTGCAGCTGAAGGAGAGCGTGGAGAT CAACTGCACCCGCCCAACAACAACACCCGCAAGAGCATCACCATCGGCCCCGGCCGCCCT TCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCCACTGCAACATCAGCGGCGAG AAGTGGAACAACACCCTGAAGCAGATCGTGACCAAGCTGCAGGCCCAGTTCGGCAACAAGAC CATCGTGTTCAAGCAGAGCAGCGGCGGCGACCCCGAGATCGTGATGCACAGCTTCAACTGCG GCGGCGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAACAGCACCTGGAACAACACCATCG CCAAGGCCATGTACGCCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGCAACATCACCGGCC TGCTGCTGACCCGCGACGGCGCAAGGAGATCAGCAACACCACCGAGATCTTCCGCCCCGGC GGCGGCGACATGCGCGACAACTGGCGCAGCGAGCTGTACAAGTACAAGTGGTGAAGATCGA GCCCTGGGCGTGGCCCCACCAAGGCCAAGCGCCGCGTGGTGCAGCGCGAGAAGCGCGCG TGACCCTGGGCGCCATGTTCCTGGGCTTCCTGGGCGCCGCCAGCACCATGGGCGCCCGCA GCCTGACCCTGACCGTGCAGGCCCGCCAGCTGCTGAGCGGCATCGTGCAGCAGCAGAACAAC CTGCTGCGCGCCATCGAGGCCCAGCAGCACCTGCTGCAGCTGACCGTGTGGGGCATCAAGCA GCTGCAGGCCCGCGTGCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTGGGCATCT GGGGCTGCAGCGGCAAGCTGATCTGCACCACCGCCGTGCCCTGGAACGCCAGCTGGAGCAAC CTACACCAACCTGATCTACACCCTGATCGAGGAGGAGCCAGAACCAGCAGGAGAAGAACGAGC AGGAGCTGCTGGAGCTGGACAAGTGGGCCAGCCTGTGGAACTGGTTCGACATCAGCAAGTGG CTGTGGTACATCAAGATCTTCATCATGATCGTGGGCGGCCTGGTGGGCCTGCGCATCGTGTTC ACCGTGCTGAGCATCGTGAACCGCGTGCGCCAGGGCTACAGCCCCCTGAGCTTCCAGACCCGC CCGCGACCGCAGCACCCCTGGTGCACGGCCTGCTGGCCCTGATCTGGGACGACCTGCGCAG GCTGCTGGGCCGCCGCGGCTGGAGGCCCTGAAGTACTGGGGCAACCTGCTGCAGTACTGGA TCCAGGAGCTGAAGAACAGCGCCGTGAGCCTGTTCGACGCCATCGCCATCGCCGTGGCCGAG GGCACCGACCGCATCATCGAGGTGGCCCAGCGCATCGGCCGCGCCTTCCTGCACATCCCCCGC CGCATCCGCCAGGGCTTCGAGCGCGCCCTGCTGTAACTCGAG

#### **SEQ ID NO:14 ILE424-ALA433**

GTCTTCGTTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTG TGGAAGGAGGCCACCACCACCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT GCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCCCAGGAGATCGTGCT GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGAAGCTGACCCCCCTGTGCGTG ACCCTGCACTGCACCAACCTGAAGAACGCCACCAACACCAAGAGCAGCAACTGGAAGGAGAT GGACCGCGGGAGATCAAGAACTGCAGCTTCAAGGTGACCACCAGCATCCGCAACAAGATGC AGAAGGAGTACGCCCTGTTCTACAAGCTGGACGTGGTGCCCATCGACAACGACAACACCAGC TACAAGCTGATCAACTGCAACACCAGCGTGATCACCCAGGCCTGCCCCAAGGTGAGCTTCGA GCCCATCCCATCCACTACTGCGCCCCGCCGGCTTCGCCATCCTGAAGTGCAACGACAAGAA GTTCAACGGCAGCGCCCCTGCACCAACGTGAGCACCGTGCAGTGCACCCACGGCATCCGCC CCGTGGTGAGCACCCAGCTGCTGAACGGCAGCCTGGCCGAGGAGGGCGTGGTGATCCGC AGCGAGAACTTCACCGACAACGCCAAGACCATCATCGTGCAGCTGAAGGAGAGCGTGGAGAT CAACTGCACCCGCCCAACAACAACACCCGCAAGAGCATCACCATCGGCCCCGGCCGCCCT TCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCCACTGCAACATCAGCGGCGAG AAGTGGAACAACACCCTGAAGCAGATCGTGACCAAGCTGCAGGCCCAGTTCGGCAACAAGAC CATCGTGTTCAAGCAGAGCAGCGGCGGCGACCCCGAGATCGTGATGCACAGCTTCAACTGCG GCGGCGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAACAGCACCTGGAACAACACCATCG GCCCCAACACCCAACGGCACCATCACCCTGCCCTGCCGCATCAAGCAGATCATCGGCGGC GCCATGTACGCCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGCAACATCACCGGCCTGCTG CTGACCCGCGACGCGCAAGGAGATCAGCAACACCACCGAGATCTTCCGCCCCGGCGGCGG CGACATGCGCGACAACTGGCGCAGCGAGCTGTACAAGTACAAGGTGGTGAAGATCGAGCCCC TGGGCGTGGCCCCACCAAGGCCAAGCGCCGCGTGGTGCAGCGCGAGAAGCGCGCGTGACC CTGGGCGCCATGTTCCTGGGCTTCCTGGGCGCCGCGGCAGCACCATGGGCGCCCGCAGCCTG ACCCTGACCGTGCAGGCCCGCCAGCTGCTGAGCGCATCGTGCAGCAGCAGCAGCAACCTGCT GCGCGCCATCGAGGCCCAGCAGCACCTGCTGCAGCTGACCGTGTGGGGCATCAAGCAGCTGC AGGCCCGCGTGCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTGGGCATCTGGGGC TGCAGCGGCAAGCTGATCTGCACCACCGCCGTGCCCTGGAACGCCAGCTGGAGCAACAAGAG CCTGGACCAGATCTGGAACAACATGACCTGGATGGAGTGGGAGCGCGAGATCGACAACTACA CCAACCTGATCTACACCCTGATCGAGGAGAGCCAGAACCAGCAGGAGAAGAACGAGCAGGA GCTGCTGGAGCTGGACAAGTGGGCCAGCCTGTGGAACTGGTTCGACATCAGCAAGTGGCTGT GGTACATCAAGATCTTCATCATGATCGTGGGCGGCCTGGTGGGCCTGCGCATCGTGTTCACCG TGCTGAGCATCGTGAACCGCGTGCGCCAGGGCTACAGCCCCCTGAGCTTCCAGACCCGCTTCC CCGCCCCGCGGCCCGACCGCCCGAGGGCATCGAGGAGGAGGGCGGCGAGCGCGACCGC GACCGCAGCAGCCCCTGGTGCACGGCCTGCTGGCCCTGATCTGGGACGACCTGCGCAGCCTG CTGGGCCGCCGCGGCTGGGAGGCCCTGAAGTACTGGGGCAACCTGCTGCAGTACTGGATCCA GGAGCTGAAGAACAGCGCCGTGAGCCTGTTCGACGCCATCGCCATCGCCGTGGCCGAGGGCA CCGACCGCATCATCGAGGTGGCCCAGCGCATCGGCCGCGCCTTCCTGCACATCCCCCGCCGCA TCCGCCAGGGCTTCGAGCGCGCCCTGCTGTAACTCGAG

## **SEQ ID NO:15 ILE423-MET434**

GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGTGTGGAGCA GTCTTCGTTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTG TGGAAGGAGGCCACCACCACCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT GCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCCCAACCCCCAGGAGATCGTGCT GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGAAGCTGACCCCCCTGTGCGTG ACCCTGCACTGCACCAACCTGAAGAACGCCACCAACACCAAGAGCAGCAACTGGAAGGAGAT GGACCGCGGCGAGATCAAGAACTGCAGCTTCAAGGTGACCACCAGCATCCGCAACAAGATGC AGAAGGAGTACGCCTGTTCTACAAGCTGGACGTGGTGCCCATCGACAACGACAACACCAGC TACAAGCTGATCAACTGCAACACCAGCGTGATCACCCAGGCCTGCCCCAAGGTGAGCTTCGA GCCCATCCCATCCACTACTGCGCCCCGCCGGCTTCGCCATCCTGAAGTGCAACGACAAGAA GTTCAACGGCAGCGCCCCTGCACCAACGTGAGCACCGTGCAGTGCACCCACGGCATCCGCC CCGTGGTGAGCACCCAGCTGCTGCACGGCAGCCTGGCCGAGGAGGGCGTGGTGATCCGC AGCGAGAACTTCACCGACAACGCCAAGACCATCATCGTGCAGCTGAAGGAGAGCGTGGAGAT CAACTGCACCGCCCAACAACAACACCCGCAAGAGCATCACCATCGGCCCCGGCCGCCCT TCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCCACTGCAACATCAGCGGCGAG AAGTGGAACAACACCCTGAAGCAGATCGTGACCAAGCTGCAGGCCCAGTTCGGCAACAAGAC CATCGTGTTCAAGCAGAGCAGCGGCGGCGACCCCGAGATCGTGATGCACAGCTTCAACTGCG GCGGCGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAACAGCACCTGGAACACCACCATCG TACGCCCCCCATCCGCGCCAGATCCGCTGCAGCAGCAACATCACCGGCCTGCTGACC CGCGACGGCGGCAAGGAGATCAGCAACACCACCGAGATCTTCCGCCCCGGCGGCGGCGACAT GCGCGACAACTGGCGCAGCGAGCTGTACAAGTACAAGTGGTGAAGATCGAGCCCCTGGGCG TGGCCCCACCAAGGCCAAGCGCGCGTGGTGCAGCGCGAGAAGCGCGCGTGACCCTGGGC GCCATGTTCCTGGGCTTCCTGGGCGCCGCCGCAGCATGGGCGCCCGCAGCCTGACCCTG ACCGTGCAGGCCCGCCAGCTGCTGAGCGGCATCGTGCAGCAGCAGAACAACCTGCTGCGCGC CATCGAGGCCCAGCAGCACCTGCTGCAGCTGACCGTGTGGGGCATCAAGCAGCTGCAGGCCC GCGTGCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTGGGCATCTGGGGCTGCAGC GGCAAGCTGATCTGCACCACCGCCGTGCCCTGGAACGCCAGCTGGAGCAACAAGAGCCTGGA CCAGATCTGGAACAACATGACCTGGATGGAGTGGGAGCGCGAGATCGACAACTACACCAACC TGATCTACACCCTGATCGAGGAGAGCCAGAACCAGCAGGAGAAGAACGAGCAGGAGCTGCTG GAGCTGGACAAGTGGGCCAGCCTGTGGAACTGGTTCGACATCAGCAAGTGGCTGTGGTACAT CAAGATCTTCATCATGATCGTGGGCGGCCTGGTGGGCCTGCGCATCGTGTTCACCGTGCTGAG CATCGTGAACCGCGTGCGCCAGGGCTACAGCCCCCTGAGCTTCCAGACCCGCTTCCCCGCCC AGCAGCCCCTGGTGCACGGCCTGCTGGCCCTGATCTGGGACGACCTGCGCAGCCTGTGCCTG CGCCGCGGCTGGGAGGCCCTGAAGTACTGGGGCAACCTGCTGCAGTACTGGATCCAGGAGCT GAAGAACAGCGCCGTGAGCCTGTTCGACGCCATCGCCATCGCCGTGGCCGAGGGCACCGACC GCATCATCGAGGTGGCCCAGCGCATCGGCCGCCTTCCTGCACATCCCCGCCGCATCCGCC AGGGCTTCGAGCGCGCCCTGCTGTAACTCGAG

## **SEQ ID NO:16 GLN422-TYR435**

GTCTTCGTTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTG TGGAAGGAGGCCACCACCACCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT GCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCCCAGGAGATCGTGCT GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGAAGCTGACCCCCCTGTGCGTG ACCCTGCACTGCACCAACCTGAAGAACGCCACCAACACCAAGAGCAGCAACTGGAAGGAGAT GGACCGCGGCGAGATCAAGAACTGCAGCTTCAAGGTGACCACCAGCATCCGCAACAAGATGC AGAAGGAGTACGCCCTGTTCTACAAGCTGGACGTGGTGCCCATCGACAACGACAACACCAGC TACAAGCTGATCAACTGCAACACCAGCGTGATCACCCAGGCCTGCCCCAAGGTGAGCTTCGA GCCCATCCCATCCACTACTGCGCCCCGCCGGCTTCGCCATCCTGAAGTGCAACGACAAGAA GTTCAACGCAGCGCCCCTGCACCAACGTGAGCACCGTGCAGTGCACCCACGGCATCCGCC CCGTGGTGAGCACCCAGCTGCTGCAACGGCAGCCTGGCCGAGGAGGGCGTGGTGATCCGC AGCGAGAACTTCACCGACAACGCCAAGACCATCATCGTGCAGCTGAAGGAGAGCGTGGAGAT CAACTGCACCCGCCCAACAACACCCCGCAAGAGCATCACCATCGGCCCCGGCCGCGCCT TCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCCACTGCAACATCAGCGGCGAG AAGTGGAACAACACCCTGAAGCAGATCGTGACCAAGCTGCAGGCCCAGTTCGGCAACAAGAC CATCGTGTTCAAGCAGAGCAGCGGCGGCGACCCCGAGATCGTGATGCACAGCTTCAACTGCG GCGGCGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAACAGCACCTGGAACAACACCATCG GCCCCAACACACCAACGGCACCATCACCCTGCCCTGCCGCATCAAGCAGGGCGGCTACGCC CCCCCATCCGCGCCAGATCCGCTGCAGCAGCAACATCACCGGCCTGCTGCTGACCCGCGAC GGCGGCAAGGAGATCAGCAACACCACCGAGATCTTCCGCCCCGGCGGCGGCGACATGCGCGA CAACTGGCGCAGCGAGCTGTACAAGTACAAGGTGGTGAAGATCGAGCCCCTGGGCGTGGCCC CCACCAAGGCCAAGCGCGCGTGGTGCAGCGCGAGAAGCGCGCGTGACCCTGGGCGCCATG TTCCTGGGCTTCCTGGGCGCCGCCGGCAGCACCATGGGCGCCCGCAGCCTGACCCTG CAGGCCCGCCAGCTGCTGAGCGCATCGTGCAGCAGCAGCAACCTGCTGCGCGCCATCGA GGCCCAGCAGCACCTGCTGCAGCTGACCGTGTGGGGCATCAAGCAGCTGCAGGCCCGCGTGC TGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTGGGCATCTGGGGCTGCAGCGGCAAG CTGATCTGCACCACCGCCGTGCCCTGGAACGCCAGCTGGAGCAACAAGAGCCTGGACCAGAT CTGGAACAACATGACCTGGATGGAGTGGGAGCCGAGATCGACAACTACACCAACCTGATCT ACACCCTGATCGAGGAGAGCCAGAACCAGCAGGAGAACGAGCAGGAGCTGCTGGAGCT GGACAAGTGGGCCAGCCTGTGGAACTGGTTCGACATCAGCAAGTGGCTGTGGTACATCAAGA TCTTCATCATGATCGTGGGCGGCCTGGTGGGCCTGCGCATCGTGTTCACCGTGCTGAGCATCG TGAACCGCGTGCGCCAGGGCTACAGCCCCCTGAGCTTCCAGACCCGCTTCCCCGCCCCCCGCG CCCCTGGTGCACGGCCTGCTGGCCCTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTCAG CGGCTGGGAGGCCCTGAAGTACTGGGGCAACCTGCTGCAGTACTGGATCCAGGAGCTGAAGA ATCGAGGTGGCCCAGCGCATCGGCCGCGCCTTCCTGCACATCCCCCGCCGCATCCGCCAGGGC TTCGAGCGCCCCTGCTGTAACTCGAG

## SEQ ID NO:17 GLN422-TYR435B

GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGTGTGGAGCA GTCTTCGTTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTG TGGAAGGAGGCCACCACCACCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT GCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCCCAGGAGATCGTGCT GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGAAGCTGACCCCCCTGTGCGTG ACCCTGCACTGCACCAACCTGAAGAACGCCACCAACACCAAGAGCAGCAACTGGAAGGAGAT GGACCGCGGCGAGATCAAGAACTGCAGCTTCAAGGTGACCACCAGCATCCGCAACAAGATGC AGAAGGAGTACGCCCTGTTCTACAAGCTGGACGTGGTGCCCATCGACAACGACAACACCAGC TACAAGCTGATCAACTGCAACACCAGCGTGATCACCCAGGCCTGCCCCAAGGTGAGCTTCGA GCCCATCCCATCCACTACTGCGCCCCGCCGGCTTCGCCATCCTGAAGTGCAACGACAAGAA GTTCAACGCCAGCGCCCCTGCACCAACGTGAGCACCGTGCAGTGCACCCACGGCATCCGCC CCGTGGTGAGCACCCAGCTGCTGCAACGGCAGCCTGGCCGAGGAGGGCGTGGTGATCCGC AGCGAGAACTTCACCGACAACGCCAAGACCATCATCGTGCAGCTGAAGGAGAGCGTGGAGAT CAACTGCACCCGCCCAACAACAACACCCGCAAGAGCATCACCATCGGCCCCGGCCGCCCCT TCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCCACTGCAACATCAGCGGCGAG AAGTGGAACACCCTGAAGCAGATCGTGACCAAGCTGCAGGCCCAGTTCGGCAACAAGAC CATCGTGTTCAAGCAGAGCAGCGGCGGCGACCCCGAGATCGTGATGCACAGCTTCAACTGCG GCGGCGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAACAGCACCTGGAACAACACCATCG GCCCCAACACCCAACGGCACCATCACCCTGCCCTGCCGCATCAAGCAGGCCCCCTACGCCC CCCCATCCGCGCCAGATCCGCTGCAGCAGCAACATCACCGGCCTGCTGCTGACCCGCGACG GCGGCAAGGAGATCAGCAACACCACCGAGATCTTCCGCCCGGCGGCGGCGACATGCGCGAC AACTGGCGCAGCGAGCTGTACAAGTACAAGGTGGTGAAGATCGAGCCCCTGGGCGTGGCCCC CACCAAGGCCAAGCGCGCGTGGTGCAGCGCGAGAAGCGCGCCGTGACCCTGGGCGCCATGT TCCTGGGCTTCCTGGGCGCCGCCGGCAGCACCATGGGCGCCCGCAGCCTGACCGTGC AGGCCCGCCAGCTGCTGAGCGCCATCGTGCAGCAGCAGCACCACCTGCTGCGCGCCATCGAG GCCCAGCAGCACCTGCTGCAGCTGACCGTGTGGGGCATCAAGCAGCTGCAGGCCCGCGTGCT GGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTGGGGCATCTGGGGCTGCAGCGGCAAGC TGATCTGCACCACCGCCGTGCCCTGGAACGCCAGCTGGAGCAACAAGAGCCTGGACCAGATC TGGAACAACATGACCTGGATGGAGTGGGAGCGCGAGATCGACAACTACACCAACCTGATCTA CACCTGATCGAGGAGAGCCAGAACCAGCAGGAGAACAGAGCAGGAGCTGCTGGAGCTG GACAAGTGGCCAGCCTGTGGAACTGGTTCGACATCAGCAAGTGGCTGTGGTACATCAAGAT CTTCATCATGATCGTGGGCGGCCTGGGCCTGCGCATCGTGTTCACCGTGCTGAGCATCGT GAACCGCGTGCGCCAGGGCTACAGCCCCTGAGCTTCCAGACCCGCTTCCCCGCCCCCCGCGG CCCGACCGCCCGAGGGCATCGAGGAGGAGGGCGGCGACCGCGACCGCAGCAGC CCCTGGTGCACGCCTGCTGGCCCTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTCAGC TACCACCGCCTGCGCGACCTGATCCTGATCGCCGCCCGCATCGTGGAGCTGCTGGGCCGCCGC GGCTGGAGCCCTGAAGTACTGGGGCAACCTGCTGCAGTACTGGATCCAGGAGCTGAAGAA CGAGGTGGCCCAGCGCATCGGCCGCGCCTTCCTGCACATCCCCCGCCGCATCCGCCAGGGCTT CGAGCGCCCTGCTGTAACTCGAG

## SEQ ID NO:18: LEU122-SER199; ARG426-GLY431

GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGTGTGGAGCA GTCTTCGTTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTG TGGAAGGAGGCCACCACCACCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT GCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCCCAGGAGATCGTGCT GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGAAGCTGGGCAACAGCGTGAT CACCCAGGCCTGCCCCAAGGTGAGCTTCGAGCCCATCCCCATCCACTACTGCGCCCCGCCGG CTTCGCCATCCTGAAGTGCAACGACAAGAAGTTCAACGGCAGCGGCCCCTGCACCAACGTGA GCACCGTGCAGTGCACCCACGGCATCCGCCCCGTGGTGAGCACCCAGCTGCTGCAACGGC AGCCTGGCCGAGGAGGCGTGGTGATCCGCAGCGAGAACTTCACCGACAACGCCAAGACCAT CATCGTGCAGCTGAAGGAGAGCGTGGAGATCAACTGCACCCGCCCCAACAACAACACCCGCA AGAGCATCACCATCGGCCCGGCCGCCCTTCTACGCCACCGGCGACATCATCGGCGACATCC GCCAGGCCCACTGCAACATCAGCGGCGAGAAGTGGAACAACACCCTGAAGCAGATCGTGACC AAGCTGCAGGCCCAGTTCGGCAACAAGACCATCGTGTTCAAGCAGAGCAGCGGCGGCGACCC CGAGATCGTGATGCACAGCTTCAACTGCGGCGGCGAGTTCTTCTACTGCAACAGCACCCAGCT GTTCAACAGCACCTGGAACAACACCATCGGCCCCAACAACACCCAACGGCACCATCACCCTGC CCTGCCGCATCAAGCAGATCATCAACCGCGGCGGCGAAGGCCATGTACGCCCCCCCATCC GCGGCCAGATCCGCTGCAGCAGCAACATCACCGGCCTGCTGCTGACCCGCGACGGCGGCAAG GAGATCAGCAACACCACCGAGATCTTCCGCCCCGGCGGCGGCGACACTGCGCGACAACTGGCG CAGCGAGCTGTACAAGTACAAGGTGGTGAAGATCGAGCCCCTGGGCGTGGCCCCACCAAGG CCAAGCGCCGCGTGGTGCAGCGCGAGAAGCGCGCCGTGACCCTGGGCGCCATGTTCCTGGGC TTCCTGGGCGCCGCCGCAGCACCATGGGCGCCCGCAGCCTGACCGTGCAGGCCCGC CAGCTGCTGAGCGCATCGTGCAGCAGCAGCACCACCTGCTGCGCGCCATCGAGGCCCAGCA GCACCTGCTGCAGCTGACCGTGTGGGGCATCAAGCAGCTGCAGGCCCGCGTGCTGGCCGTGG AGCGCTACCTGAAGGACCAGCAGCTGCTGGGCATCTGGGGCTGCAGCGGCAAGCTGATCTGC ACCACCGCCGTGCCCTGGAACGCCAGCTGGAGCAACAAGAGCCTGGACCAGATCTGGAACAA CATGACCTGGATGGAGTGGGAGCCCGAGATCGACAACTACACCAACCTGATCTACACCCTGA TCGAGGAGAGCCAGAACCAGCAGGAGAAGAACGAGCAGGAGCTGCTGGAGCTGGACAAGTG GGCCAGCCTGTGGAACTGGTTCGACATCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCAT GATCGTGGGCGGCCTGGGGCCTGCGCATCGTGTTCACCGTGCTGAGCATCGTGAACCGCGT GCGCCAGGGCTACAGCCCCCTGAGCTTCCAGACCCGCTTCCCCGCCCCCCGCGGCCCCGACCG CCCCGAGGCATCGAGGAGGAGGGCGGCGAGCGCGACCGCAGCAGCCCCCTGGTGC ACGGCCTGCTGGCCCTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTCAGCTACCACCGCC TGCGCGACCTGATCCTGATCGCCGCCCGCATCGTGGAGCTGCTGGGCCGCCGCGGCTGGGAGG CCCTGAAGTACTGGGGCAACCTGCTGCAGTACTGGATCCAGGAGCTGAAGAACAGCGCCGTG CAGCGCATCGGCCGCCTTCCTGCACATCCCCCGCCGCATCCGCCAGGGCTTCGAGCGCGCC **CTGCTGTAACTCGAG** 

#### SEQ ID NO:19 LEU122-SER199; ARG426-LYS432

GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGTGTGGAGCA GTCTTCGTTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTG TGGAAGGAGGCCACCACCACCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT GCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCCCAGGAGATCGTGCT GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGAAGCTGGGCAACAGCGTGAT CACCCAGGCCTGCCCCAAGGTGAGCTTCGAGCCCATCCCATCCACTACTGCGCCCCGCCGG CTTCGCCATCCTGAAGTGCAACGACAAGAAGTTCAACGGCAGCGGCCCCTGCACCAACGTGA GCACCGTGCAGTGCACCCACGGCATCCGCCCCGTGGTGAGCACCCAGCTGCTGCAGCGC AGCCTGGCCGAGGAGGGCGTGGTGATCCGCAGCGAGACTTCACCGACAACGCCAAGACCAT CATCGTGCAGCTGAAGGAGAGCGTGGAGATCAACTGCACCCGCCCAACAACACCCCGCA AGAGCATCACCATCGGCCCGGCCGCGCTTCTACGCCACCGGCGACATCATCGGCGACATCC GCCAGGCCCACTGCAACATCAGCGGCGAGAAGTGGAACAACACCCTGAAGCAGATCGTGACC AAGCTGCAGGCCCAGTTCGGCAACAAGACCATCGTGTTCAAGCAGAGCAGCGGCGGCGACCC CGAGATCGTGATGCACAGCTTCAACTGCGGCGGCGAGTTCTTCTACTGCAACAGCACCCAGCT GTTCAACAGCACCTGGAACAACACCATCGGCCCCAACAACACCAACGGCACCATCACCCTGC CCTGCCGCATCAAGCAGATCATCAACCGCGGCGGCAACAAGGCCATGTACGCCCCCCCATCC GCGCCAGATCCGCTGCAGCAGCAACATCACCGGCCTGCTGCTGACCCGCGACGGCGGCAAG GAGATCAGCAACACCACCGAGATCTTCCGCCCCGGCGGCGGCGACACTGCGCGACAACTGGCG CAGCGAGCTGTACAAGTACAAGGTGGTGAAGATCGAGCCCCTGGGCGTGGCCCCCACCAAGG CCAAGCGCCGCGTGGTGCAGCGCGAGAAGCGCGCGTGACCCTGGGCGCCATGTTCCTGGGC TTCCTGGGCGCCGGCAGCACCATGGGCGCCCGCAGCCTGACCCTGACCGTGCAGGCCCGC CAGCTGCTGAGCGCATCGTGCAGCAGCAGAACAACCTGCTGCGCGCCATCGAGGCCCAGCA GCACCTGCTGCAGCTGACCGTGTGGGGCATCAAGCAGCTGCAGGCCCGCGTGCTGGCCGTGG AGCGCTACCTGAAGGACCAGCAGCTGCTGGGCATCTGGGGCTGCAGCGGCAAGCTGATCTGC ACCACCGCCGTGCCCTGGAACGCCAGCTGGAGCAACAAGAGCCTGGACCAGATCTGGAACAA CATGACCTGGATGGAGTGGGAGCGCGAGATCGACAACTACACCAACCTGATCTACACCCTGA TCGAGGAGACCAGAACCAGCAGGAGAAGAACGAGCAGGAGCTGCTGGAGCTGGACAAGTG GGCCAGCCTGTGGAACTGGTTCGACATCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCAT GATCGTGGGCCTGGTGGGCCTGCGCATCGTGTTCACCGTGCTGAGCATCGTGAACCGCGT GCGCCAGGGCTACAGCCCCTGAGCTTCCAGACCGCTTCCCCGCCCCCGCGGCCCGACCG ACGGCCTGCTGGCCCTGATCTGGGACGACCTGCGCAGCCTGTTCCAGCTACCACCGCC TGCGCGACCTGATCCTGATCGCCGCCCGCATCGTGGAGCTGCTGGGCCGCCGCGCTGGGAGG CCCTGAAGTACTGGGGCAACCTGCTGCAGTACTGGATCCAGGAGCTGAAGAACAGCGCCGTG CAGCGCATCGGCCGCCTTCCTGCACATCCCCCGCCGCATCCGCCAGGGCTTCGAGCGCGCC CTGCTGTAACTCGAG

#### SEO ID NO: 20: LEU122-SER199; TRP427-GLY431

GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGTGTGGAGCA GTCTTCGTTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTG TGGAAGGAGGCCACCACCACCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT GCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCCAGGAGATCGTGCT GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGAAGCTGGGCAACAGCGTGAT CACCCAGGCCTGCCCAAGGTGAGCTTCGAGCCCATCCCCATCACTACTGCGCCCCGCCGG CTTCGCCATCCTGAAGTGCAACGACAAGAAGTTCAACGGCAGCGGCCCCTGCACCAACGTGA GCACCGTGCAGTGCACCCACGGCATCCGCCCCGTGGTGAGCACCCAGCTGCTGCAACGGC AGCCTGGCCGAGGGGGGGGTGATCCGCAGCGAGAACTTCACCGACAACGCCAAGACCAT CATCGTGCAGCTGAAGGAGAGCGTGGAGATCAACTGCACCCGCCCCAACAACAACACCCGCA AGAGCATCACCATCGGCCCGGCCGCCTTCTACGCCACCGGCGACATCATCGGCGACATCC GCCAGGCCCACTGCAACATCAGCGGCGAGAAGTGGAACACACCCTGAAGCAGATCGTGACC AAGCTGCAGGCCCAGTTCGGCAACAAGACCATCGTGTTCAAGCAGAGCAGCGGCGGCGACCC CGAGATCGTGATGCACAGCTTCAACTGCGGCGGCGAGTTCTTCTACTGCAACAGCACCCAGCT GTTCAACAGCACCTGGAACAACACCATCGGCCCCAACAACACCAACGGCACCATCACCCTGC CCTGCCGCATCAAGCAGATCATCAACCGCTGGGGCGGCAAGGCCATGTACGCCCCCCCATCC GCGCCAGATCCGCTGCAGCAGCAACATCACCGGCCTGCTGACCCGCGACGGCGACAG CAGCGAGCTGTACAAGTACAAGGTGGTGAAGATCGAGCCCCTGGGCGTGGCCCCCACCAAGG CCAAGCGCCGCGTGCGCGCGCGAGAAGCGCGCGTGACCCTGGGCGCCATGTTCCTGGGC TTCCTGGGCGCCGCCGCAGCACCATGGGCGCCCGCAGCCTGACCCTGACCGTGCAGGCCCGC CAGCTGCTGAGCGGCATCGTGCAGCAGCAGCACCACCTGCTGCGCGCCATCGAGGCCCAGCA GCACCTGCTGCAGCTGTGGGGCATCAAGCAGCTGCAGGCCCGCGTGCTGGCCGTGG AGCGCTACCTGAAGGACCAGCAGCTGCTGGGCCATCTGGGGCTGCAGCGGCAAGCTGATCTGC ACCACCGCCGTGCCCTGGAACGCCAGCTGGAGCAACAAGAGCCTGGACCAGATCTGGAACAA CATGACCTGGATGGAGTGGGAGCGCGAGATCGACAACTACACCAACCTGATCTACACCCTGA TCGAGGAGAGCCAGAACCAGCAGGAGAAGAACGAGCAGGAGCTGCTGGACCAAGTG GGCCAGCCTGTGGAACTGGTTCGACATCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCAT GATCGTGGGCCTGGTGGGCCTGCGCATCGTGTTCACCGTGCTGAGCATCGTGAACCGCGT GCGCCAGGGCTACAGCCCCCTGAGCTTCCAGACCGGCTTCCCCGCCCCCCGGGGCCCCGACCG CCCCGAGGGCATCGAGGAGGAGGGCGGCGAGCGCGACCGCAGCAGCCCCCTGGTGC ACGGCCTGCTGGCCCTGATCTGGGACGACCTGCGCAGCCTGTTGCCTGTTCAGCTACCACCGCC TGCGCGACCTGATCCTGATCGCCGCCCGCATCGTGGAGCTGCTGGGCCGCCGCGGCTGGGAGG CCCTGAAGTACTGGGGCAACCTGCTGCAGTACTGGATCCAGGAGCTGAAGAACAGCGCCGTG AGCCTGTTCGACGCCATCGCCATCGCCGTGGCCGAGGGCACCGCATCATCGAGGTGGCC CAGCGCATCGGCCGCCTTCCTGCACATCCCCCGCCGCATCCGCCAGGGCTTCGAGCGCGCC **CTGCTGTAACTCGAG** 

# SEQ ID NO:21 LYS121-VAL200; ASN425-LYS432

GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGTGTGGAGCA GTCTTCGTTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTG TGGAAGGAGGCCACCACCACCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT GCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCCCAGGAGATCGTGCT GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACACATGGTGGAGCAGATGCACGAG GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGAAGGCCCCCGTGATCACCCA GGCCTGCCCCAAGGTGAGCTTCGAGCCCATCCCCATCCACTACTGCGCCCCGCCGGCTTCGC CATCCTGAAGTGCAACGACAAGAAGTTCAACGGCAGCGGCCCCTGCACCAACGTGAGCACCG TGCAGTGCACCCACGGCATCCGCCCCGTGGTGAGCACCCAGCTGCTGCTGAACGGCAGCCTGG CCGAGGAGGCGTGGTGATCCGCAGCGAGAACTTCACCGACAACGCCAAGACCATCATCGTG CAGCTGAAGGAGAGCGTGGAGATCAACTGCACCCGCCCAACAACAACACCCGCAAGAGCAT CACCATCGGCCCGGCCGCCCTTCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGC CCACTGCAACATCAGCGGCGAGAAGTGGAACAACACCCTGAAGCAGATCGTGACCAAGCTGC AGGCCCAGTTCGGCAACAAGACCATCGTGTTCAAGCAGAGCAGCGGCGGCGACCCCGAGATC GTGATGCACAGCTTCAACTGCGGCGGCGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAAC AGCACCTGGAACAACACCATCGGCCCCAACAACACCAACGGCACCATCACCCTGCCCTGCCG CATCAAGCAGATCATCAACGCCCCCAAGGCCATGTACGCCCCCCCATCCGCGGCCAGATCCG CTGCAGCAGCAACATCACCGGCCTGCTGCTGACCCGCGACGGCGAAGGAGATCAGCAACA CCACCGAGATCTTCCGCCCCGGCGGCGGCGACATGCGCGAACTGGCGCAGCGAGCTGTAC AAGTACAAGGTGGTGAAGATCGAGCCCCTGGGCGTGGCCCCACCAAGGCCAAGCGCCGCGT GGTGCAGCGCGAGAAGCGCGCCGTGACCCTGGGCGCCATGTTCCTGGGCTTCCTGGGCGCCGC CGGCAGCACCATGGGCGCCCGCAGCCTGACCCTGACCGTGCAGGCCCGCCAGCTGCTGAGCG GCATCGTGCAGCAGCAGCAGCACCTGCTGCGCGCCCATCGAGGCCCCAGCAGCACCTGCTGCAG CTGACCGTGTGGGGCATCAAGCAGCTGCAGGCCCGCGTGCTGGCCGTGGAGCGCTACCTGAA GGACCAGCAGCTGCTGGGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCGCCGTGC CCTGGAACGCCAGCTGGAGCAACAAGAGCCTGGACCAGATCTGGAACAACATGACCTGGATG GAGTGGGAGCGCGAGATCGACAACTACACCAACCTGATCTACACCCTGATCGAGGAGAGCCA GAACCAGCAGGAGAAGAACGAGCAGGAGCTGCTGGAGCTGGACAAGTGGGCCAGCCTGTGG AACTGGTTCGACATCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCATGATCGTGGGCGGC CTGGTGGGCCTGCGCATCGTGTTCACCGTGCTGAGCATCGTGAACCGCGTGCGCCAGGGCTAC AGCCCCTGAGCTTCCAGACCCGCTTCCCCGCCCCCGGGCCCGGCCCCGAGGGCATC GAGGAGGAGGCGCGACCGCGACCGCAGCAGCCCCCTGGTGCACGGCCTGCTGGC CCTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTCAGCTACCACCGCCTGCGCGACCTGAT CCTGATCGCCGCCCGCATCGTGGAGCTGCTGGGCCGCCGCGGCTGGGAGGCCCTGAAGTACTG GGGCAACCTGCTGCAGTACTGGATCCAGGAGCTGAAGAACAGCGCCGTGAGCCTGTTCGACG CCATCGCCATCGCCGTGGCCGAGGGCACCGACCGCATCGTCGAGGTGGCCCAGCGCATCGGC CGCGCCTTCCTGCACATCCCCCGCCGCATCCGCCAGGGCTTCGAGCGCGCCCTGCTGTAACTC GAG

## SEO ID NO:22 VAL120-ILE201; ILE 424-ALA433

GAATTCGCCACCATGGATGCAATGAAGAGGGGCTCTGCTGTGTGCTGCTGTGTGGAGCA GTCTTCGTTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTG TGGAAGGAGGCCACCACCACCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT GCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCCCAGGAGATCGTGCT GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGGGCGGCATCACCCAGGCCTG GAAGTGCAACGACAAGAAGTTCAACGGCAGCGGCCCCTGCACCAACGTGAGCACCGTGCAGT GCACCCACGGCATCCGCCCCGTGGTGAGCACCCAGCTGCTGCTGAACGGCAGCCTGGCCGAG GAGGGCGTGGTGATCCGCAGCGAGAACTTCACCGACAACGCCAAGACCATCATCGTGCAGCT GAAGGAGAGCGTGGAGATCAACTGCACCCGCCCCAACAACAACACCCGCAAGAGCATCACCA TCGGCCCGGCCGCCTTCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCCACT GCAACATCAGCGGCGAGAAGTGGAACAACACCCTGAAGCAGATCGTGACCAAGCTGCAGGCC CAGTTCGGCAACAAGACCATCGTGTTCAAGCAGAGCAGCGGCGGCGACCCCGAGATCGTGAT GCACAGCTTCAACTGCGGCGGGGGGTTCTTCTACTGCAACAGCACCCAGCTGTTCAACAGCAC CTGGAACAACACCATCGGCCCCAACAACACCAACGGCACCATCACCCTGCCCTGCCGCATCA AGCAGATCATCGGCGGCCCTGTACGCCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGC AACATCACCGGCCTGCTGCTGACCCGCGACGGCGCAAGGAGATCAGCAACACCACCGAGAT CTTCCGCCCGGCGGCGCGACATGCGCGACAACTGGCGCAGCGAGCTGTACAAGTACAAGG TGGTGAAGATCGAGCCCCTGGGCGTGGCCCCACCAAGGCCAAGCGCCGCGTGGTGCAGCGC GAGAAGCGCCCTGACCCTGGGCGCCATGTTCCTGGGCTTCCTGGGCGCCCCGCCAGCACC ATGGGCGCCGCAGCCTGACCCTGACCGTGCAGGCCCGCCAGCTGCTGAGCGGCATCGTGCA GCAGCAGAACAACCTGCTGCGCGCCATCGAGGCCCAGCAGCACCTGCTGCAGCTGACCGTGT GGGGCATCAAGCAGCTGCAGGCCCGCGTGCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAG CTGCTGGGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCGCCGTGCCCTGGAACGCC CGAGATCGACAACTACACCAACCTGATCTACACCCTGATCGAGGAGAGCCAGAACCAGCAGG AGAAGAACGAGCAGGAGCTGCTGGAGCTGGACAAGTGGGCCAGCCTGTGGAACTGGTTCGAC ATCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCATGATCGTGGGCGGCCTGGTGGGCCTG CGCATCGTGTTCACCGTGCTGAGCATCGTGAACCGCGTGCGCCAGGGCTACAGCCCCCTGAGC TTCCAGACCCGCTTCCCCGCCCCCGCGGCCCCGACCGCCCCGAGGGCATCGAGGAGGAGGG CGGCGAGCGCGACCGCAGCAGCCCCCTGGTGCACGGCCTGCTGGCCCTGATCTGGG ACGACCTGCGCAGCCTGTGCCTGTTCAGCTACCACCGCCTGCGCGACCTGATCCTGATCGCCG CCCGCATCGTGGAGCTGCTGGGCCGCCGCGGCTGGGAGGCCCTGAAGTACTGGGGCAACCTG CTGCAGTACTGGATCCAGGAGCTGAAGAACAGCGCCGTGAGCCTGTTCGACGCCATCGCCATC GCCGTGGCCGAGGGCACCGCATCATCGAGGTGGCCCAGCGCATCGGCCGCCCTTCCT GCACATCCCCGCCGCATCCGCCAGGGCTTCGAGCGCGCCCTGCTGTAACTCGAG

#### 62 / 65

### SEQ ID NO:23: VAL120-ILE201B; ILE424-ALA433

GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGTGTGGAGCA GTCTTCGTTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTG TGGAAGGAGGCCACCACCACCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT GCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCCCAGGAGATCGTGCT GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGCCCGGCATCACCCAGGCCTGC CCCAAGGTGAGCTTCGAGCCCATCCCATCCACTACTGCGCCCCGCCGGCTTCGCCATCCTG AAGTGCAACGACAAGAAGTTCAACGGCAGCGGCCCCTGCACCAACGTGAGCACCGTGCAGTG CACCCACGGCATCCGCCCGTGGTGAGCACCCAGCTGCTGAACGGCAGCCTGGCCGAGG AGGGCGTGGTGATCCGCAGCGAGAACTTCACCGACAACGCCAAGACCATCATCGTGCAGCTG AAGGAGAGCGTGGAGATCAACTGCACCCGCCCCAACAACAACACCCCGCAAGAGCATCACCAT CGGCCCCGGCCGCCTTCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCCACTG CAACATCAGCGGCGAGAAGTGGAACAACACCCTGAAGCAGATCGTGACCAAGCTGCAGGCCC AGTTCGGCAACAAGACCATCGTGTTCAAGCAGAGCAGCGGCGGCGACCCCGAGATCGTGATG CACAGCTTCAACTGCGGCGGCGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAACAGCACC TGGAACAACACCATCGGCCCCAACACACCACGGCACCATCACCCTGCCCTGCCGCATCAA GCAGATCATCGGCGGCGCCATGTACGCCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGCA ACATCACCGGCCTGCTGACCCGCGACGGCGGCAAGGAGATCAGCAACACCACCGAGATC TTCCGCCCGGCGGCGGCGACATGCGCGACACTGGCGCAGCGAGCTGTACAAGTACAAGGT GGTGAAGATCGAGCCCCTGGGCGTGGCCCCCACCAAGGCCAAGCGCCGCGTGGTGCAGCGCG AGAAGCGCGCCGTGACCCTGGGCGCCATGTTCCTGGGCTTCCTGGGCGCCGCCGGCAGCACCA TGGGCGCCCGCAGCCTGACCCTGACCGTGCAGGCCCGCCAGCTGCTGAGCGGCATCGTGCAG CAGCAGAACAACCTGCTGCGCGCCATCGAGGCCCAGCAGCACCTGCTGCAGCTGACCGTGTG GGGCATCAAGCAGCTGCAGGCCCGCGTGCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGC TGCTGGGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCGCCGTGCCCTGGAACGCCA GAGATCGACAACTACACCAACCTGATCTACACCCTGATCGAGGAGAGCCAGAACCAGCAGGA GAAGAACGAGCAGGAGCTGCTGGAGCTGGACAAGTGGGCCAGCCTGTGGAACTGGTTCGACA GCATCGTGTTCACCGTGCTGAGCATCGTGAACCGCGTGCGCCAGGGCTACAGCCCCCTGAGCT TCCAGACCCGCTTCCCCGCCCCCGCGGCCCCGACCGCCCGAGGGCATCGAGGAGGAGGGC GGCGAGCGCGACCGCAGCAGCCCCCTGGTGCACGGCCTGCTGGCCCTGATCTGGGA CGACCTGCGCAGCCTGTTCCGCTACCACCGCCTGCGCGACCTGATCCTGATCGCCGC CCGCATCGTGGAGCTGCTGGGCCGCCGCGGCTGGAGGCCCTGAAGTACTGGGGCAACCTGC TGCAGTACTGGATCCAGGAGCTGAAGAACAGCGCCGTGAGCCTGTTCGACGCCATCGCCATC GCCGTGGCCGAGGGCACCGACCGCATCATCGAGGTGGCCCAGCGCATCGGCCGCCTTCCT GCACATCCCCGCCGCATCCGCCAGGGCTTCGAGCGCGCCCTGCTGTAACTCGAG

# SEQ ID NO:24 VAL120-THR202; ILE424-ALA433

GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGTGTGGAGCA GTCTTCGTTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTG TGGAAGGAGGCCACCACCACCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT GCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCCCAGGAGATCGTGCT GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGGGCGGCGCCACCCAGGCCTG GAAGTGCAACGACAAGAAGTTCAACGGCAGCGGCCCCTGCACCAACGTGAGCACCGTGCAGT GCACCCACGCATCCGCCCCGTGGTGAGCACCCAGCTGCTGAACGGCAGCCTGGCCGAG GAGGGCGTGGTGATCCGCAGCGAGAACTTCACCGACAACGCCAAGACCATCATCGTGCAGCT GAAGGAGAGCGTGGAGATCAACTGCACCCGCCCAACAACAACACCCGCAAGAGCATCACCA TCGGCCCCGGCCGCCTTCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCCACT GCAACATCAGCGGCGAGAAGTGGAACAACACCCTGAAGCAGATCGTGACCAAGCTGCAGGCC CAGTTCGGCAACAAGACCATCGTGTTCAAGCAGAGCAGCGGCGGCGACCCCGAGATCGTGAT GCACAGCTTCAACTGCGGCGGCGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAACAGCAC CTGGAACAACACCATCGGCCCCAACAACACCAACGGCACCATCACCCTGCCCTGCCGCATCA AGCAGATCATCGGCGGCGCCATGTACGCCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGC AACATCACCGGCCTGCTGACCCGCGACGGCGGCAAGGAGATCAGCAACACCACCGAGAT CTTCCGCCCCGGCGCGCGACATGCGCGACAACTGGCGCAGCGAGCTGTACAAGTACAAGG TGGTGAAGATCGAGCCCCTGGGCGTGGCCCCCACCAAGGCCCAAGCGCCGCGTGGTGCAGCGC GAGAAGCGCGCCGTGACCCTGGGCGCCATGTTCCTGGGCTTCCTGGGCGCCGCCGGCAGCACC ATGGGCGCCCGCAGCCTGACCCTGACCGTGCAGGCCCCAGCTGCTGAGCGGCATCGTGCA GCAGCAGAACAACCTGCTGCGCGCCATCGAGGCCCAGCAGCACCTGCTGCAGCTGACCGTGT GGGGCATCAAGCAGCTGCAGGCCCGCGTGCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAG CTGCTGGGCATCTGGGGCTGCAGCGCAAGCTGATCTGCACCACCGCCGTGCCCTGGAACGCC CGAGATCGACAACTACACCAACCTGATCTACACCCTGATCGAGGAGAGCCAGAACCAGCAGG AGAAGAACGAGCAGGAGCTGCTGGAGCTGGACAAGTGGGCCAGCCTGTGGAACTGGTTCGAC ATCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCATGATCGTGGGCGGCCTGGTGGGCCTG CGCATCGTGTTCACCGTGCTGAGCATCGTGAACCGCGTGCGCCAGGGCTACAGCCCCCTGAGC TTCCAGACCCGCTTCCCCGCCCCCGCGGCCCCGACCGCCCCGAGGGCATCGAGGAGGAGGG ACGACCTGCGCAGCCTGTGCCTGTTCAGCTACCACCGCCTGCGCGACCTGATCCTGATCGCCG CTGCAGTACTGGATCCAGGAGCTGAAGAACAGCGCCGTGAGCCTGTTCGACGCCATCGCCATC GCCGTGGCCGAGGGCACCGACCGCATCATCGAGGTGGCCCAGCGCATCGGCCGCCCTTCCT GCACATCCCCGCGCATCCGCCAGGGCTTCGAGCGCGCCCTGCTGTAACTCGAG

#### **SEQ ID NO:25 VAL127-ASN195**

GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGTGTGGAGCA GTCTTCGTTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTG TGGAAGGAGGCCACCACCACCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT GCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCCCAGGAGATCGTGCT GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGTGAAGCTGACCCCCTGTGCGTG GGGGCAGGGAACTGCAACACCAGCGTGATCACCCAGGCCTGCCCCAAGGTGAGCTTCGAGCC CATCCCCATCCACTACTGCGCCCCGCCGGCTTCGCCATCCTGAAGTGCAACGACAAGAAGTT CAACGCAGCGCCCCTGCACCAACGTGAGCACCGTGCAGTGCACCCACGGCATCCGCCCG TGGTGAGCACCCAGCTGCTGCAACGGCAGCCTGGCCGAGGAGGGCGTGGTGATCCGCAGC GAGAACTTCACCGACAACGCCAAGACCATCATCGTGCAGCTGAAGGAGAGCGTGGAGATCAA CTGCACCCGCCCAACAACACCCCGCAAGAGCATCACCATCGGCCCCGGCCGCCCTTCTA CGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCCACTGCAACATCAGCGGCGAGAAGT GGAACAACACCCTGAAGCAGATCGTGACCAAGCTGCAGGCCCAGTTCGGCAACAAGACCATC GTGTTCAAGCAGAGCAGCGGCGCGACCCCGAGATCGTGATGCACAGCTTCAACTGCGGCGG CGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAACAGCACCTGGAACAACACCATCGGCCC CAACAACACCAACGGCACCATCACCCTGCCCTGCCGCATCAAGCAGATCATCAACCGCTGGC AGGAGGTGGGCAAGGCCATGTACGCCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGCAAC ATCACCGGCCTGCTGACCCGCGACGGCGACGGCAAGGAGATCAGCAACACCACCGAGATCTT CCGCCCGGCGGCGCGACATGCGCGACAACTGGCGCAGCGAGCTGTACAAGTACAAGGTGG TGAAGATCGAGCCCTGGGCGTGGCCCCACCAAGGCCAAGCGCCGCGTGGTGCAGCGCGAG AAGCGCGCGTGACCCTGGGCGCCATGTTCCTGGGCTTCCTGGGCGCCGCCGGCAGCACCATG GGCGCCGCAGCCTGACCCTGACCGTGCAGGCCCGCCAGCTGCTGAGCGGCATCGTGCAGCA GCAGAACAACCTGCTGCGCGCCATCGAGGCCCAGCAGCACCTGCTGCAGCTGACCGTGTGGG GCATCAAGCAGCTGCAGGCCCGCGTGCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTG CTGGGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCGCCGTGCCCTGGAACGCCAG AGATCGACAACTACACCAACCTGATCTACACCCTGATCGAGGAGAGCCAGAACCAGCAGGAG AAGAACGAGCAGGAGCTGCTGGAGCTGGACAAGTGGGCCAGCCTGTGGAACTGGTTCGACAT CAGCAAGTGGCTGTGGTACATCAAGATCTTCATCATGATCGTGGGCGGCCTGGTGGGCCTGCG CATCGTGTTCACCGTGCTGAGCATCGTGAACCGCGTGCGCCAGGGCTACAGCCCCCTGAGCTT CCAGACCCGCTTCCCCGCCCCCGCGGCCCCGACCGCCCCGAGGGCATCGAGGAGGAGGGCG GCGAGCGCGACCGCAGCAGCAGCCCCTGGTGCACGGCCTGCTGGCCCTGATCTGGGAC GACCTGCGCAGCCTGTTCCAGCTACCACCGCCTGCGCGACCTGATCCTGATCGCCGCC CGCATCGTGGAGCTGCTGGGCCGCCGCGGCTGGGAGGCCCTGAAGTACTGGGGCAACCTGCT GCAGTACTGGATCCAGGAGCTGAAGAACAGCGCCGTGAGCCTGTTCGACGCCATCGCCATCG CCGTGGCCGAGGGCACCGCATCATCGAGGTGGCCCAGCGCATCGGCCGCGCTTCCTGC ACATCCCCGCCGCATCCGCCAGGGCTTCGAGCGCGCCCTGCTGTAACTCGAG

# SEQ ID NO:26 VAL127-ASN195; ARG426-GLY431

GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGTGTGGAGCA GTCTTCGTTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTG TGGAAGGAGCCACCACCACCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT GCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCCCAGGAGATCGTGCT GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGAAGCTGACCCCCCTGTGCGTG GGGGCAGGGAACTGCAACACCAGCGTGATCACCCAGGCCTGCCCCAAGGTGAGCTTCGAGCC CATCCCCATCCACTACTGCGCCCCGCCGGCTTCGCCATCCTGAAGTGCAACGACAAGAAGTT CAACGGCAGCGCCCCTGCACCAACGTGAGCACCGTGCAGTGCACCCACGGCATCCGCCCCG TGGTGAGCACCCAGCTGCTGAACGGCAGCCTGGCCGAGGAGGGCGTGGTGATCCGCAGC GAGAACTTCACCGACAACGCCAAGACCATCATCGTGCAGCTGAAGGAGAGCGTGGAGATCAA CTGCACCCGCCCAACAACAACACCCGCAAGAGCATCACCATCGGCCCCGGCCGCCCTTCTA CGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCCACTGCAACATCAGCGGCGAGAAGT GGAACAACACCCTGAAGCAGATCGTGACCAAGCTGCAGGCCCAGTTCGGCAACAAGACCATC GTGTTCAAGCAGAGCAGCGGCGACCCCGAGATCGTGATGCACAGCTTCAACTGCGGCGG CGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAACAGCACCTGGAACAACACCATCGGCCC CAACAACACCAACGGCACCATCACCCTGCCCTGCCGCATCAAGCAGATCATCAACCGCGGCG GCGGCAAGGCCATGTACGCCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGCAACATCACC GGCCTGCTGACCCGCGACGGCGGCAAGGAGATCAGCAACACCACCGAGATCTTCCGCCC CGGGGGCGGCGACATGCGCGACAACTGGCGCAGCGAGCTGTACAAGTACAAGTGGTGAAG ATCGAGCCCCTGGGCGTGGCCCCACCAAGGCCAAGCGCCGCGTGGTGCAGCGCGAGAAGCG CGCCGTGACCCTGGGCGCATGTTCCTGGGCTTCCTGGGCGCCGCCGGCAGCACCATGGGCGC CCGCAGCCTGACCCTGCAGGCCCGCCAGCTGCTGAGCGGCATCGTGCAGCAGCAGA ACAACCTGCTGCGCGCCATCGAGGCCCAGCAGCACCTGCTGCAGCTGACCGTGTGGGGCATC AAGCAGCTGCAGGCCCGCGTGCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTGGG CATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCGCCGTGCCCTGGAACGCCAGCTGGA GACAACTACACCAACCTGATCTACACCCTGATCGAGGAGAGCCAGAACCAGCAGGAGAAGAA CGAGCAGGAGCTGCTGGACCAAGTGGGCCAGCCTGTGGAACTGGTTCGACATCAGCA TGTTCACCGTGCTGAGCATCGTGAACCGCGTGCGCCAGGGCTACAGCCCCCTGAGCTTCCAGA CGCGACCGCGACCGCAGCCCCCTGGTGCACGGCCTGCTGGCCCTGATCTGGGACGACCTG GTGGAGCTGCTGGGCCGCCGCGGCTGGGAGGCCCTGAAGTACTGGGGCAACCTGCTGCAGTA CTGGATCCAGGAGCTGAAGAACAGCGCCGTGAGCCTGTTCGACGCCATCGCCATCGCCGTGG  ${\tt CCGAGGGCACCGACCGCATCATCGAGGTGGCCCAGCGCATCGGCCGCGCCTTCCTGCACATCC}$ CCCGCCGCATCCGCCAGGGCTTCGAGCGCCCCTGCTGTAACTCGAG

### SEQUENCE LISTING

- <110> Chiron Corporation
- <120> MODIFIED HIV ENV POLYPEPTIDES
- <130> 1605.100
- <140>
- <141>
- <160> 26
- <170> PatentIn Ver. 2.0
- <210> 1
- <211> 856
- <212> PRT
- <213> Human immunodeficiency virus
- <400> 1
- Met Arg Val Lys Glu Lys Tyr Gln His Leu Trp Arg Trp Gly Trp Arg
  1 5 10 15
- Trp Gly Thr Met Leu Leu Gly Met Leu Met Ile Cys Ser Ala Thr Glu 20 25 30
- Lys Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala 35 40 45
- Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Asp Thr Glu 50 60
- Val His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn 65 70 75 80
- Pro Gln Glu Val Val Leu Val Asn Val Thr Glu Asn Phe Asn Met Trp
  85 90 95
- Lys Asn Asp Met Val Glu Gln Met His Glu Asp Ile Ile Ser Leu Trp
  100 105 110
- Asp Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Ser 115 120 125
- Leu Lys Cys Thr Asp Leu Lys Asn Asp Thr Asn Thr Asn Ser Ser Ser 130 135 140
- Gly Arg Met Ile Met Glu Lys Gly Glu Ile Lys Asn Cys Ser Phe Asn 145 150 155 160
- Ile Ser Thr Ser Ile Arg Gly Lys Val Gln Lys Glu Tyr Ala Phe Phe 165 170 175
- Tyr Lys Leu Asp Ile Ile Pro Ile Asp Asn Asp Thr Thr Ser Tyr Lys 180 185 190

Leu Thr Ser Cys Asn Thr Ser Val Ile Thr Gln Ala Cys Pro Lys Val

- Ser Phe Glu Pro Ile Pro Ile His Tyr Cys Ala Pro Ala Gly Phe Ala 210 215 220
- Ile Leu Lys Cys Asn Asn Lys Thr Phe Asn Gly Thr Gly Pro Cys Thr 225 230 235 240
- Asn Val Ser Thr Val Gln Cys Thr His Gly Ile Arg Pro Val Val Ser 245 250 255
- Thr Gln Leu Leu Asn Gly Ser Leu Ala Glu Glu Glu Val Val Ile
  260 265 270
- Arg Ser Val Asn Phe Thr Asp Asn Ala Lys Thr Ile Ile Val Gln Leu 275 280 285
- Asn Thr Ser Val Glu Ile Asn Cys Thr Arg Pro Asn Asn Asn Thr Arg 290 295 300
- Lys Arg Ile Arg Ile Gln Arg Gly Pro Gly Arg Ala Phe Val Thr Ile 305 310 315 320
- Gly Lys Ile Gly Asn Met Arg Gln Ala His Cys Asn Ile Ser Arg Ala 325 330 335
- Lys Trp Asn Asn Thr Leu Lys Gln Ile Ala Ser Lys Leu Arg Glu Gln \$340\$ \$350\$
- Phe Gly Asn Asn Lys Thr Ile Ile Phe Lys Gln Ser Ser Gly Gly Asp 355 360 365
- Pro Glu Ile Val Thr His Ser Phe Asn Cys Gly Glu Phe Phe Tyr 370 375 380
- Cys Asn Ser Thr Gln Leu Phe Asn Ser Thr Trp Phe Asn Ser Thr Trp 385 390 395 400
- Ser Thr Glu Gly Ser Asn Asn Thr Glu Gly Ser Asp Thr Ile Thr Leu 405 410 415
- Pro Cys Arg Ile Lys Gln Ile Ile Asn Met Trp Gln Lys Val Gly Lys 420 425 430
- Ala Met Tyr Ala Pro Pro Ile Ser Gly Gln Ile Arg Cys Ser Ser Asn 435 440 445
- Ile Thr Gly Leu Leu Leu Thr Arg Asp Gly Gly Asn Ser Asn Asn Glu 450 460
- Ser Glu Ile Phe Arg Pro Gly Gly Gly Asp Met Arg Asp Asn Trp Arg 465 470 475 480
- Ser Glu Leu Tyr Lys Tyr Lys Val Val Lys Ile Glu Pro Leu Gly Val
  485 490 495
- Ala Pro Thr Lys Ala Lys Arg Arg Val Val Gln Arg Glu Lys Arg Ala 500 505 510

Val Gly Ile Gly Ala Leu Phe Leu Gly Phe Leu Gly Ala Ala Gly Ser 515 520 525

- Thr Met Gly Ala Ala Ser Met Thr Leu Thr Val Gln Ala Arg Gln Leu 530 540
- Leu Ser Gly Ile Val Gln Gln Gln Asn Asn Leu Leu Arg Ala Ile Glu 545 550 555 560
- Ala Gln Gln His Leu Leu Gln Leu Thr Val Trp Gly Ile Lys Gln Leu
  565 570 575
- Gln Ala Arg Ile Leu Ala Val Glu Arg Tyr Leu Lys Asp Gln Gln Leu 580 585 590
- Leu Gly Ile Trp Gly Cys Ser Gly Lys Leu Ile Cys Thr Thr Ala Val
- Pro Trp Asn Ala Ser Trp Ser Asn Lys Ser Leu Glu Gln Ile Trp Asn 610 615 620
- His Thr Thr Trp Met Glu Trp Asp Arg Glu Ile Asn Asn Tyr Thr Ser 625 630 635 640
- Leu Ile His Ser Leu Ile Glu Glu Ser Gln Asn Gln Glu Lys Asn 645 650 655
- Glu Gln Glu Leu Glu Leu Asp Lys Trp Ala Ser Leu Trp Asn Trp
  660 665 670
- Phe Asn Ile Thr Asn Trp Leu Trp Tyr Ile Lys Leu Phe Ile Met Ile 675 680 685
- Val Gly Gly Leu Val Gly Leu Arg Ile Val Phe Ala Val Leu Ser Ile 690 695 700
- Val Asn Arg Val Arg Gln Gly Tyr Ser Pro Leu Ser Phe Gln Thr His 705 710 715 720
- Leu Pro Thr Pro Arg Gly Pro Asp Arg Pro Glu Gly Ile Glu Glu Glu 725 730 735
- Gly Gly Glu Arg Asp Arg Asp Arg Ser Ile Arg Leu Val Asn Gly Ser 740 745 750
- Leu Ala Leu Ile Trp Asp Asp Leu Arg Ser Leu Cys Leu Phe Ser Tyr 755 760 765
- His Arg Leu Arg Asp Leu Leu Leu Ile Val Thr Arg Ile Val Glu Leu
  770 780
- Leu Gly Arg Arg Gly Trp Glu Ala Leu Lys Tyr Trp Trp Asn Leu Leu 785 790 795 800
- Gln Tyr Trp Ser Gln Glu Leu Lys Asn Ser Ala Val Ser Leu Leu Asn 805 810 815
- Ala Thr Ala Ile Ala Val Ala Glu Gly Thr Asp Arg Val Ile Glu Val 820 825 830

Val Gln Gly Ala Cys Arg Ala Ile Arg His Ile Pro Arg Arg Ile Arg 835 840 845

Gln Gly Leu Glu Arg Ile Leu Leu 850 855

<210> 2

<211> 847

<212> PRT

<213> Human immunodeficiency virus

<400> 2

Met Arg Val Lys Gly Ile Arg Lys Asn Tyr Gln His Leu Trp Arg Gly
1 5 10 15

Gly Thr Leu Leu Gly Met Leu Met Ile Cys Ser Ala Val Glu Lys  $20 \hspace{1cm} 25 \hspace{1cm} 30$ 

Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala Thr 35 40 45

Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Asp Thr Glu Val

His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro 65 70 75 80

Gln Glu Ile Val Leu Glu Asn Val Thr Glu Asn Phe Asn Met Trp Lys 85 90 95

Asn Asn Met Val Glu Gln Met His Glu Asp Ile Ile Ser Leu Trp Asp 100 105 110

Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu 115 120 125

His Cys Thr Asn Leu Lys Asn Ala Thr Asn Thr Lys Ser Ser Asn Trp

Lys Glu Met Asp Arg Gly Glu Ile Lys Asn Cys Ser Phe Lys Val Thr 145 150 155 160

Thr Ser Ile Arg Asn Lys Met Gln Lys Glu Tyr Ala Leu Phe Tyr Lys 165 170 175

Leu Asp Val Val Pro Ile Asp Asn Asp Asn Thr Ser Tyr Lys Leu Ile 180 185 190

Asn Cys Asn Thr Ser Val Ile Thr Gln Ala Cys Pro Lys Val Ser Phe 195 200 205

Glu Pro Ile Pro Ile His Tyr Cys Ala Pro Ala Gly Phe Ala Ile Leu 210 215 220

Lys Cys Asn Asp Lys Lys Phe Asn Gly Ser Gly Pro Cys Thr Asn Val 225 230 235 240

 Ser
 Thr
 Val
 Gln
 Cys
 Thr
 His
 Gly
 Ile
 Arg
 Pro
 Val
 Val
 Ser
 Thr
 Gln

 Leu
 Leu
 Leu
 Asn
 Gly
 Ser
 Leu
 Ala
 Glu
 Gly
 Val
 Val
 Ile
 Arg
 Ser

 Glu
 Asn
 Phe
 Thr
 Asp
 Asn
 Ala
 Lys
 Thr
 Ile
 Ile
 Ile
 Val
 Gln
 Leu
 Lys
 Glu
 Asn
 Thr
 Asn
 Asn
 Asn
 Thr
 Asn
 Asn

Trp Gln Glu Val Gly Lys Ala Met Tyr Ala Pro Pro Ile Arg Gly Gln 420 425 430

Ile Arg Cys Ser Ser Asn Ile Thr Gly Leu Leu Leu Thr Arg Asp Gly 435 440 445

Gly Lys Glu Ile Ser Asn Thr Thr Glu Ile Phe Arg Pro Gly Gly Gly 450 455 460

Asp Met Arg Asp Asn Trp Arg Ser Glu Leu Tyr Lys Tyr Lys Val Val 465 470 475 480

Lys Ile Glu Pro Leu Gly Val Ala Pro Thr Lys Ala Lys Arg Arg Val
485 490 495

Val Gln Arg Glu Lys Arg Ala Val Thr Leu Gly Ala Met Phe Leu Gly 500 505 510

Phe Leu Gly Ala Ala Gly Ser Thr Met Gly Ala Arg Ser Leu Thr Leu 515 520 525

Thr Val Gln Ala Arg Gln Leu Leu Ser Gly Ile Val Gln Gln Gln Asn 530 535 540

Asn Leu Leu Arg Ala Ile Glu Ala Gln Gln His Leu Leu Gln Leu Thr 545 550 555 560

Val Trp Gly Ile Lys Gln Leu Gln Ala Arg Val Leu Ala Val Glu Arg 565 570 575

Tyr Leu Lys Asp Gln Gln Leu Leu Gly Ile Trp Gly Cys Ser Gly Lys 580 585 590

Leu Ile Cys Thr Thr Ala Val Pro Trp Asn Ala Ser Trp Ser Asn Lys 595 600 605

Ser Leu Asp Gln Ile Trp Asn Asn Met Thr Trp Met Glu Trp Glu Arg 610 615 620

Glu Ile Asp Asn Tyr Thr Asn Leu Ile Tyr Thr Leu Ile Glu Glu Ser 625 635 640

Gln Asn Gln Glu Lys Asn Glu Gln Glu Leu Leu Glu Leu Asp Lys 645 650 655

Trp Ala Ser Leu Trp Asn Trp Phe Asp Ile Ser Lys Trp Leu Trp Tyr
660 665 670

Ile Lys Ile Phe Ile Met Ile Val Gly Gly Leu Val Gly Leu Arg Ile 675 680 685

Val Phe Thr Val Leu Ser Ile Val Asn Arg Val Arg Gln Gly Tyr Ser 690 695 700

Pro Leu Ser Phe Gln Thr Arg Phe Pro Ala Pro Arg Gly Pro Asp Arg 705 710 715 720

Pro Glu Gly Ile Glu Glu Glu Gly Gly Glu Arg Asp Arg Asp Arg Ser 725 730 735

Ser Pro Leu Val His Gly Leu Leu Ala Leu Ile Trp Asp Asp Leu Arg 740 745 750

Ser Leu Cys Leu Phe Ser Tyr His Arg Leu Arg Asp Leu Ile 11e 755 760 765

Ala Ala Arg Ile Val Glu Leu Leu Gly Arg Arg Gly Trp Glu Ala Leu 770 775 780

Lys Tyr Trp Gly Asn Leu Leu Gln Tyr Trp Ile Gln Glu Leu Lys Asn 785 790 795 800

Ser Ala Val Ser Leu Phe Asp Ala Ile Ala Ile Ala Val Ala Glu Gly 805 810 815

Thr Asp Arg Ile Ile Glu Val Ala Gln Arg Ile Gly Arg Ala Phe Leu 820 825 830

His Ile Pro Arg Arg Ile Arg Gln Gly Phe Glu Arg Ala Leu Leu 835 840 845

<210> 3

<211> 2310

<212> DNA

<213> Artificial Sequence

<220> <223> Description of Artificial Sequence: Val120-Ala204 gaattegeea ceatggatge aatgaagaga gggetetget gtgtgetget getgtqtqqa 60 gcagtcttcg tttcgcccag cgccgtggag aagctgtggg tgaccgtgta ctacggcgtg 120 cccgtgtgga aggaggccac caccaccctg ttctgcgcca gcgacgccaa ggcctacgac 180 accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 240 gagatcgtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtggag 300 cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgggcgcc 360 ggegeetgee ceaaggtgag ettegageee atceceatee actaetgege eeeeggeegge 420 ttcgccatcc tgaagtgcaa cgacaagaag ttcaacggca gcggcccctg caccaacgtg 480 agcaccgtgc agtgcaccca cggcatccgc cccgtggtga gcacccagct gctgctgaac 540 ggcagcctgg ccgaggaggg cgtggtgatc cgcagcgaga acttcaccga caacgccaag 600 accatcatcg tgcagctgaa ggagagcgtg gagatcaact gcacccgccc caacaacaac 660 accegeaaga geateaceat eggeeegge egegeettet aegeeacegg egacateate 720 ggcgacatcc gccaggccca ctgcaacatc agcggcgaga agtggaacaa caccctgaag 780 cagategtga ccaagetgca ggcccagtte ggcaacaaga ccategtgtt caageagage 840 ageggeggeg acceegagat egtgatgeac agetteaact geggeggega gttettetac 900 tgcaacagca cccagctgtt caacagcacc tggaacaaca ccatcggccc caacaacacc 960 aacggcacca tcaccctgcc ctgccgcatc aagcagatca tcaaccgctg gcaggaggtg 1020 ggcaaggcca tgtacgcccc ccccatccgc ggccagatcc gctgcagcag caacatcacc 1080 ggcctgctgc tgacccgcga cggcggcaag gagatcagca acaccaccga gatcttccgc 1140 cccggcggcg gcgacatgcg cgacaactgg cgcagcgagc tgtacaagta caaggtggtg 1200 aagatcgagc ccctgggcgt ggcccccacc aaggccaagc gccgcgtggt gcagcgcgag 1260 aagcgcgccg tgaccctggg cgccatgttc ctgggcttcc tgggcgccgc cggcagcacc 1320 atgggegece geageetgae cetgaeegtg caggeeegee agetgetgag eggeategtg 1380 cagcagcaga acaacctgct gcgcgccatc gaggcccagc agcacctgct gcagctgacc 1440 gtgtggggca tcaagcagct gcaggcccgc gtgctggccg tggagcgcta cctgaaggac 1500 cagcagetgc tgggcatetg gggctgcagc ggcaagetga tetgcaccac cgccgtgccc 1560 tggaacgcca gctggagcaa caagagcctg gaccagatct ggaacaacat gacctggatg 1620 gagtgggagc gcgagatcga caactacacc aacctgatct acaccctgat cgaggagagc 1680 cagaaccagc aggagaagaa cgagcaggag ctgctggagc tggacaagtg ggccagcctg 1740 tggaactggt tcgacatcag caagtggctg tggtacatca agatcttcat catgatcgtg 1800 ggcggcctgg tgggcctgcg catcgtgttc accgtgctga gcatcgtgaa ccgcgtgcgc 1860 cagggetaca geoccetgag ettecagace egetteceeg ecceegegg eccegacege 1920 cccgagggca tcgaggagga gggcggcgag cgcgaccgcg accgcagcag ccccctggtg 1980 cacggcctgc tggccctgat ctgggacgac ctgcgcagcc tgtgcctgtt cagctaccac 2040 cgcctgcgcg acctgatect gategccgcc cgcategtgg agetgctggg ccgccgcggc 2100 tgggaggccc tgaagtactg gggcaacctg ctgcagtact ggatccagga gctgaagaac 2160 agegeegtga geetgttega egeeategee ategeegtgg eegagggeae egaeegeate 2220 ategaggtgg cocagegeat eggeegegee tteetgeaca teceeegeeg cateegeeag 2280 ggcttcgagc gcgccctgct gtaactcgag <210> 4 <211> 2316 <212> DNA <213> Artificial Sequence <220> <223> Description of Artificial Sequence: Val120-Ile201 <400> 4 gaattegeea ceatggatge aatgaagaga gggetetget gtgtgetget getgtgtgga 60 gcagtcttcg tttcgcccag cgccgtggag aagctgtggg tgaccgtgta ctacggcgtg 120 cccgtgtgga aggaggccac caccacctg ttctgcgcca gcgacgccaa ggcctacgac 180 accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 240 gagatcgtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtggag 300 cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgggcggc 360

```
atcacccagg cctgccccaa ggtgagcttc gagcccatcc ccatccacta ctgcgcccc 420
gccggcttcg ccatcctgaa gtgcaacgac aagaagttca acggcagcgg cccctgcacc 480
aacgtgagca ccgtgcagtg cacccacggc atccgccccg tggtgagcac ccagctgctg 540
ctgaacggca gcctggccga ggagggcgtg gtgatccgca gcgagaactt caccgacaac 600
gccaagacca tcatcgtgca gctgaaggag agcgtggaga tcaactgcac ccgccccaac 660
aacaacaccc gcaagagcat caccatcggc cccggccgcg ccttctacgc caccggcgac 720
atcatcggcg acatccgcca ggcccactgc aacatcagcg gcgagaagtg gaacaacacc 780
ctgaagcaga tcgtgaccaa gctgcaggcc cagttcggca acaagaccat cgtgttcaag 840
cagageageg geggegaeee egagategtg atgeaeaget teaactgegg eggegagtte 900
ttctactgca acagcaccca gctgttcaac agcacctgga acaacaccat cgqccccaac 960
aacaccaacg gcaccatcac cctgccctgc cgcatcaagc agatcatcaa ccgctggcag 1020
gaggtgggca aggccatgta cgccccccc atccgcggcc agatccgctg cagcagcaac 1080
atcaccggcc tgctgctgac ccgcgacggc ggcaaggaga tcagcaacac caccgagatc 1140
ttccgccccg gcggcggcga catgcgcgac aactggcgca gcgagctgta caagtacaag 1200
gtggtgaaga tcgagcccct gggcgtggcc cccaccaagg ccaagcgccg cgtggtgcag 1260
cgcgagaagc gcgccgtgac cctgggcgcc atgttcctgg gcttcctggg cgccgccggc 1320
agcaccatgg gcgcccgcag cctgaccctg accgtgcagg cccgccagct gctgagcggc 1380
atcgtgcagc agcagaacaa cctgctgcgc gccatcgagg cccagcagca cctgctgcag 1440
ctgaccgtgt ggggcatcaa gcagctgcag gcccgcgtgc tggccgtgga gcgctacctg 1500
aaggaccagc agctgctggg catctggggc tgcagcggca agctgatctg caccaccgcc 1560
gtgccctgga acgccagctg gagcaacaag agcctggacc agatctggaa caacatgacc 1620
tggatggagt gggagcgcga gatcgacaac tacaccaacc tgatctacac cctgatcgag 1680
gagagccaga accagcagga gaagaacgag caggagctgc tggagctgga caagtgggcc 1740
agcctgtgga actggttcga catcagcaag tggctgtggt acatcaagat cttcatcatg 1800
atcgtgggcg gcctggtggg cctgcgcate gtgttcaccg tgctgagcat cgtgaaccgc 1860
gtgcgccagg gctacagccc cctgagcttc cagacccgct tccccgcccc ccgcggcccc 1920
gaccgccccg agggcatcga ggaggagggc ggcgagcgcg accgcgaccg cagcagcccc 1980
ctggtgcacg gcctgctggc cctgatctgg gacgacctgc gcagcctgtg cctgttcagc 2040
taccacegee tgegegaeet gateetgate geegeeegea tegtggaget getgggeege 2100
cgcggctggg aggccctgaa gtactggggc aacctgctgc agtactggat ccaggagctg 2160
aagaacagcg ccgtgagcct gttcgacgcc atcgccatcg ccgtggccga gggcaccgac 2220
cgcatcatcg aggtggccca gcgcatcggc cgcgccttcc tgcacatccc ccgccgcatc 2280
cgccagggct tcgagcgcgc cctgctgtaa ctcgag
                                                                  2316
<210> 5
<211> 2322
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: Val120-Ile201B
gaattegeea ceatggatge aatgaagaga gggetetget gtgtgetget getgtgtgga 60
geagtetteg tttegeeeag egeegtggag aagetgtggg tgaeegtgta etaeggegtg 120
cccgtgtgga aggaggccac caccacctg ttctgcgcca gcgacgccaa ggcctacgac 180
accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 240
gagatcgtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtggag 300
cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgcccggc 360
atcacccagg cctgccccaa ggtgagcttc gagcccatcc ccatccacta ctgcgcccc 420
geeggetteg ceateetgaa gtgeaacgae aagaagttea aeggeagegg eeeetgeace 480
aacgtgagca ccgtgcagtg cacccacggc atccgccccg tggtgagcac ccagctgctg 540
ctgaacggca gcctggccga ggagggcgtg gtgatccgca gcgagaactt caccgacaac 600
gccaagacca tcatcgtgca gctgaaggag agcgtggaga tcaactgcac ccgccccaac 660
aacaacaccc gcaagagcat caccateggc cccggccgcg ccttctacgc caccggcgac 720
atcatcggcg acatccgcca ggcccactgc aacatcagcg gcgagaagtg gaacaacacc 780
ctgaagcaga tcgtgaccaa gctgcaggcc cagttcggca acaagaccat cgtgttcaag 840
cagagcagcg geggegacce egagategtg atgeacaget teaactgegg eggegagtte 900
ttctactgca acagcaccca gctgttcaac agcacctgga acaacaccat cggccccaac 960
```

```
aacaccaacg gcaccatcac cetgccetgc cgcatcaagc agatcatcaa cegetggcag 1020
gaggtgggca aggccatgta cgccccccc atccgcggcc agatccgctg cagcagcaac 1080
atcaccggcc tgctgctgac ccgcgacggc ggcaaggaga tcagcaacac caccgagatc 1140
ttccgccccg gcggcggcga catgcgcgac aactggcgca gcgagctgta caagtacaag 1200
gtggtgaaga tcgagcccct gggcgtggcc cccaccaagg ccaagcgccg cgtggtgcag 1260
cgcgagaagc gcgccgtgac cctgggcgcc atgttcctgg gcttcctggg cgccgccggc 1320
ageaccatgg gegecegeag cetgaecetg accgtgeagg ceegecaget getgagegge 1380
ategtgcage ageagaacaa cetgetgege gecategagg cecageagea cetgetgcag 1440
ctgaccgtgt ggggcatcaa gcagctgcag gcccgcgtgc tggccgtgga gcgctacctg 1500
aaggaccagc agctgctggg catctggggc tgcagcggca agctgatctg caccaccgcc 1560
gtgccctgga acgccagctg gagcaacaag agcctggacc agatctggaa caacatgacc 1620
tggatggagt gggagcgcga gatcgacaac tacaccaacc tgatctacac cctgatcgag 1680
gagagccaga accagcagga gaagaacgag caggagctgc tggagctgga caagtgggcc 1740
agcctgtgga actggttcga catcagcaag tggctgtggt acatcaagat cttcatcatg 1800
atogtgggcg gcctggtggg cctgcgcatc gtgttcaccg tgctgagcat cgtgaaccgc 1860
gtgcgccagg gctacagccc cctgagcttc cagacccgct tccccgcccc ccgcggcccc 1920
gaccgccccg agggcatcga ggaggagggc ggcgagcgcg accgcgaccg cagcagcccc 1980
ctggtgcacg gcctgctggc cctgatctgg gacgacctgc gcagcctgtg cctgttcagc 2040
taccaccgcc tgcgcgacct gatcctgatc gccgcccgca tcgtggagct gctgggccgc 2100
cgcggctggg aggccctgaa gtactggggc aacctgctgc agtactggat ccaggagctg 2160
aagaacageg cogtgageet gttegacgee ategecateg cogtggeega gggcacegae 2220
cgcatcatcg aggtggccca gcgcatcggc cgcgccttcc tgcacatccc ccgccgcatc 2280
cgccagggct tcgagcgcgc cctgctgtaa ctcgagcgtg ct
<210> 6
<211> 2328
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: Lys121-Val200
<400> 6
gaattcgcca ccatggatgc aatgaagaga gggctctgct gtgtgctgct gctgtgtgga 60
gcagtcttcg tttcgcccag cgccgtggag aagctgtggg tgaccgtgta ctacggcgtg 120
cccgtgtgga aggaggccac caccaccctg ttctgcgcca gcgacgccaa ggcctacgac 180
accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 240
gagatcgtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtggag 300
cagatgeacg aggacateat cageetgtgg gaccagagee tgaageeetg egtgaaggee 360
cccgtgatca cccaggcctg ccccaaggtg agcttcgagc ccatccccat ccactactgc 420
geoceogoeg gettegecat cetgaagtge aacgacaaga agtteaacgg cageggeece 480
tgcaccaacg tgagcaccgt gcagtgcacc cacggcatcc gccccgtggt gagcacccag 540
ctgctgctga acggcagcct ggccgaggag ggcgtggtga tccgcagcga gaacttcacc 600
gacaacgcca agaccatcat cgtgcagctg aaggagagcg tggagatcaa ctgcacccgc 660
cccaacaaca acaccegcaa gagcatcacc ateggeceeg geegegeett etaegecace 720
ggcgacatca tcggcgacat ccgccaggcc cactgcaaca tcagcggcga gaagtggaac 780
aacaccetga agcagategt gaccaagetg caggeecagt teggeaacaa gaccategtg 840
ttcaagcaga gcagcggcgg cgaccccgag atcgtgatgc acagcttcaa ctgcggcggc 900
gagttettet aetgeaacag cacceagetg tteaacagea eetggaacaa caccategge 960
cccaacaaca ccaacggcac catcaccctg ccctgccgca tcaagcagat catcaaccgc 1020
tggcaggagg tgggcaaggc catgtacgcc ccccccatcc geggccagat cegetgcagc 1080
agcaacatca ccggcctgct gctgacccgc gacggcggca aggagatcag caacaccacc 1140
gagatettee geeeeggegg eggegacatg egegacaaet ggegeagega getgtacaag 1200
tacaaggtgg tgaagatega geceetggge gtggeeecca ceaaggeeaa gegeegegtg 1260
gtgcagcgcg agaagcgcgc cgtgaccctg ggcgccatgt tcctgggctt cctgggcgcc 1320
gccggcagca ccatgggcgc ccgcagcctg accctgaccg tgcaggcccg ccagctgctg 1380
ageggeateg tgeageagea gaacaacetg etgegegeea tegaggeeea geageacetg 1440
ctgcagctga ccgtgtgggg catcaagcag ctgcaggccc gcgtgctggc cgtggagcgc 1500
tacctgaagg accagcagct gctgggcatc tggggctgca gcggcaagct gatctgcacc 1560
```

```
accgccgtgc cctggaacgc cagctggagc aacaagagcc tggaccagat ctggaacaac 1620
atgacctgga tggagtggga gcgcgagatc gacaactaca ccaacctgat ctacaccctg 1680
atcgaggaga gccagaacca gcaggagaag aacgagcagg agctgctgga gctggacaag 1740
tgggccagcc tgtggaactg gttcgacatc agcaagtggc tgtggtacat caagatcttc 1800
atcatgatcg tgggcggcct ggtgggcctg cgcatcgtgt tcaccgtgct gagcatcgtg 1860
aaccgcgtgc gccagggcta cagccccctg agcttccaga cccgcttccc cgccccccgc 1920
ggccccgacc gccccgaggg catcgaggag gagggcggcg agcgcgaccg cgaccgcagc 1980
agececetgg tgeaeggeet getggeeetg atetgggaeg acetgegeag cetgtgeetg 2040
ttcagctacc accecctece ceaccteatc cteatcecce cccecatcet geaectect 2100
ggccgccgcg gctgggaggc cctgaagtac tggggcaacc tgctgcagta ctggatccag 2160
gagctgaaga acagcgccgt gagcctgttc gacgccatcg ccatcgccgt ggccgagggc 2220
acegacegea teategaggt ggeecagege ateggeegeg cetteetgea cateeceege 2280
cgcatccgcc agggcttcga gcgcgccctg ctgtaactcg agcgtgct
<210> 7
<211> 2334
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: Leu122-Ser199
<400> 7
gaattegeea ceatggatge aatgaagaga gggetetget gtgtgetget getgtgtgga 60
geagtetteg thregeceag egeegtggag aagetgtggg tgacegtgta chaeggegtg 120
cccgtgtgga aggaggccac caccacctg ttctgcgcca gcgacgccaa ggcctacgac 180
accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 240
gagategtge tggagaaegt gacegagaae tteaacatgt ggaagaaeaa catggtggag 300
cagatgcacg aggacatcat cagcetgtgg gaccagagee tgaageeetg egtgaagetg 360
ggcaacagcg tgatcaccca ggcctgcccc aaggtgagct tcgagcccat ccccatccac 420
tactgcgccc ccgccggctt cgccatcctg aagtgcaacg acaagaagtt caacggcagc 480
ggcccctgca ccaacgtgag caccgtgcag tgcacccacg gcatccgccc cgtggtgagc 540
acceagetge tgetgaacgg cageetggee gaggagggeg tggtgateeg cagegagaac 600
ttcaccgaca acgccaagac catcatcgtg cagctgaagg agagcgtgga gatcaactgc 660
accegececa acaacaacae eegcaagage ateaceateg geeceggeeg egeettetae 720
gccaccggcg acatcatcgg cgacatccgc caggcccact gcaacatcag cggcgagaag 780
tggaacaaca ccctgaagca gatcgtgacc aagctgcagg cccagttcgg caacaagacc 840
atcgtgttca agcagagcag cggcggcgac cccgagatcg tgatgcacag cttcaactgc 900
ggcggcgagt tcttctactg caacagcacc cagctgttca acagcacctg gaacaacacc 960
ateggeecca acaacaccaa eggeaccate accetgeect geegcateaa geagateate 1020
aaccgctggc aggaggtggg caaggccatg tacgccccc ccatccgcgg ccagatccgc 1080
tgcagcagca acatcaccgg cctgctgctg acccgcgacg gcggcaagga gatcagcaac 1140
accaccgaga tetteegeee eggeggegge gacatgegeg acaactggeg cagegagetg 1200
tacaagtaca aggtggtgaa gatcgagccc ctgggcgtgg cccccaccaa ggccaagcgc 1260
egegtggtge agegegagaa gegegeegtg accetgggeg ceatgtteet gggetteetg 1320
ggcgccgccg gcagcaccat gggcgcccgc agcctgaccc tgaccgtgca ggcccgccag 1380
ctgctgagcg gcatcgtgca gcagcagaac aacctgctgc gcgccatcga ggcccagcag 1440
cacctgctgc agctgaccgt gtggggcatc aagcagctgc aggcccgcgt gctggccgtg 1500
gagcgctacc tgaaggacca gcagctgctg ggcatctggg gctgcagcgg caagctgatc 1560
tgcaccaccg ccgtgccctg gaacgccagc tggagcaaca agagcctgga ccagatctgg 1620
aacaacatga cctggatgga gtgggagcgc gagatcgaca actacaccaa cctgatctac 1680
accetgateg aggagageca gaaccageag gagaagaacg ageaggaget getggagetg 1740
gacaagtggg ccagcctgtg gaactggttc gacatcagca agtggctgtg gtacatcaag 1800
atcttcatca tgatcgtggg cggcctggtg ggcctgcgca tcgtgttcac cgtgctgagc 1860
atcgtgaacc gcgtgcgcca gggctacagc cccctgagct tccagacccg cttccccgcc 1920
ccccgcggcc ccgaccgccc cgagggcatc gaggaggagg gcggcgagcg cgaccgcgac 1980
cgcagcagcc ccctggtgca cggcctgctg gccctgatct gggacgacct gcgcagcctg 2040
tgcctgttca gctaccaccg cctgcgcgac ctgatcctga tcgccgcccg catcgtggag 2100
ctgctgggcc gccgcggctg ggaggccctg aagtactggg gcaacctgct gcagtactgg 2160
```

```
atccaggage tgaagaacag cgccgtgage ctgttcgacg ccatcgccat cgccgtggcc 2220
gagggcaccg accgcatcat cgaggtggcc cagcgcatcg gccgcgcctt cctgcacatc 2280
ccccgccgca tccgccaggg cttcgagcgc gccctgctgt aactcgagcg tgct
<210> 8
<211> 2316
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: Val120-Thr202
<400> 8
gaattogoca coatggatgo aatgaagaga gggototgot gtgtgotgot gotgtgtgga 60
gcagtetteg tttegeceag egcegtggag aagetgtggg tgaeegtgta etaeggegtg 120
cccgtgtgga aggaggccac caccacctg ttctgcgcca gcgacgccaa ggcctacgac 180
accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 240
gagatcgtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtggag 300
cagatgcacg aggacatcat cagcetgtgg gaccagagee tgaageeetg egtgggegge 360
gccacccagg cctgccccaa ggtgagcttc gagcccatcc ccatccacta ctgcgccccc 420
gccggcttcg ccatcctgaa gtgcaacgac aagaagttca acggcagcgg cccctgcacc 480
aacgtgagca ccgtgcagtg cacccacggc atccgccccg tggtgagcac ccagctgctg 540
ctgaacggca gcctggccga ggagggcgtg gtgatccgca gcgagaactt caccgacaac 600
gccaagacca tcatcgtgca gctgaaggag agcgtggaga tcaactgcac ccgccccaac 660
aacaacaccc gcaagagcat caccatcggc cccggccgcg ccttctacgc caccggcgac 720
atcatcggcg acatccgcca ggcccactgc aacatcagcg gcgagaagtg gaacaacacc 780
ctgaagcaga tcgtgaccaa gctgcaggcc cagttcggca acaagaccat cgtgttcaag 840
cagagcagcg gcggcgaccc cgagatcgtg atgcacagct tcaactgcgg cggcgagttc 900
ttctactgca acagcaccca gctgttcaac agcacctgga acaacaccat cggccccaac 960
aacaccaacg gcaccatcac cctgccctgc cgcatcaagc agatcatcaa ccgctggcag 1020
gaggtgggca aggccatgta cgccccccc atccgcggcc agatccgctg cagcagcaac 1080
atcaccggcc tgctgctgac ccgcgacggc ggcaaggaga tcagcaacac caccgagatc 1140
ttccgccccg gcggcggcga catgcgcgac aactggcgca gcgagctgta caagtacaag 1200
gtggtgaaga tcgagcccct gggcgtggcc cccaccaagg ccaagcgccg cgtggtgcag 1260
cgcgagaagc gcgccgtgac cctgggcgcc atgttcctgg gcttcctggg cgccgccggc 1320
agcaccatgg gcgcccgcag cctgaccctg accgtgcagg cccgccagct gctgagcggc 1380
atcgtgcagc agcagaacaa cctgctgcgc gccatcgagg cccagcagca cctgctgcag 1440
ctgaccgtgt ggggcatcaa gcagctgcag gcccgcgtgc tggccgtgga gcgctacctg 1500
aaggaccage agetgetggg catetgggge tgcageggea agetgatetg caccacegee 1560
gtgccctgga acgccagctg gagcaacaag agcctggacc agatctggaa caacatgacc 1620
tggatggagt gggagcgcga gatcgacaac tacaccaacc tgatctacac cctgatcgag 1680
gagagccaga accagcagga gaagaacgag caggagctgc tggagctgga caagtgggcc 1740
agcctgtgga actggttcga catcagcaag tggctgtggt acatcaagat cttcatcatg 1800
ategtgggeg geetggtggg cetgegeate gtgtteaceg tgetgageat egtgaacege 1860
gtgcgccagg gctacagccc cctgagcttc cagacccgct tccccgcccc ccgcggcccc 1920
gaccgccccg agggcatcga ggaggagggc ggcgagcgcg accgcgaccg cagcagcccc 1980
ctggtgcacg gcctgctggc cctgatctgg gacgacctgc gcagcctgtg cctgttcagc 2040
taccacegee tgegegaeet gateetgate geegeeegea tegtggaget getgggeege 2100
cgcggctggg aggccctgaa gtactggggc aacctgctgc agtactggat ccaggagctg 2160
aagaacagcg ccgtgagcct gttcgacgcc atcgccatcg ccgtggccga gggcaccgac 2220
cgcatcatcg aggtggccca gcgcatcggc cgcgccttcc tgcacatccc ccgccgcatc 2280
cgccagggct tcgagcgcgc cctgctgtaa ctcgag
                                                                  2316
<210> 9
<211> 2541
<212> DNA
<213> Artificial Sequence
<220>
```

<223> Description of Artificial Sequence: Trp427-Gly431

```
<400> 9
gaattegeea ceatggatge aatgaagaga gggetetget gtgtgetget getgtgtgga 60
gcagtcttcg tttcgcccag cgccgtggag aagctgtggg tgaccgtgta ctacggcgtg 120
cccgtgtgga aggaggccac caccacctg ttctgcgcca gcgacgccaa ggcctacgac 180
accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 240
gagategtge tggagaaegt gacegagaae tteaacatgt ggaagaacaa catggtggag 300
cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaagctg 360
acceccetgt gegtgaccet geactgeace aacctgaaga acgceaceaa caccaagage 420
agcaactgga aggagatgga ccgcggcgag atcaagaact gcagcttcaa ggtgaccacc 480
agcateegea acaagatgea gaaggagtae geeetgttet acaagetgga egtggtgeee 540
ategacaacg acaacaccag ctacaagetg atcaactgca acaccagegt gatcacccag 600
gcctgcccca aggtgagctt cgagcccatc cccatccact actgcgcccc cgccggcttc 660
gccatcctga agtgcaacga caagaagttc aacggcagcg gcccctgcac caacgtgagc 720
acceptgcagt gcacccacgg catccgcccc gtggtgagca cccagctgct gctgaacggc 780
agcetggeeg aggagggegt ggtgateege agegagaact teacegacaa egecaagaee 840
atcatcgtgc agctgaagga gagcgtggag atcaactgca cccgccccaa caacaacacc 900
cgcaagagca tcaccatcgg ccccggccgc gccttctacg ccaccggcga catcatcggc 960
gacatcegee aggeeeactg caacateage ggegagaagt ggaacaacae cetgaageag 1020
atcgtgacca agctgcaggc ccagttcggc aacaagacca tcgtgttcaa gcagagcagc 1080
ggcggcgacc ccgagatcgt gatgcacagc ttcaactgcg gcggcgagtt cttctactgc 1140
aacagcaccc agctgttcaa cagcacctgg aacaacacca tcggccccaa caacaccaac 1200
ggcaccatca ccctgccctg ccgcatcaag cagatcatca accgctgggg cggcaaggcc 1260
atgtacgccc ccccatccg cggccagatc cgctgcagca gcaacatcac cggcctgctg 1320
ctgacccgcg acggcggcaa ggagatcagc aacaccaccg agatcttccg ccccggcggc 1380
ggcgacatgc gcgacaactg gcgcagcgag ctgtacaagt acaaggtggt gaagatcgag 1440
cccctgggcg tggccccac caaggccaag cgccgcgtgg tgcagcgcga gaagcgcgcc 1500
gtgaccetgg gegecatgtt cetgggette etgggegeeg eeggeageac catgggegee 1560
cgcagcctga ccctgaccgt gcaggcccgc cagctgctga gcggcatcgt gcagcagcag 1620
aacaacctgc tgcgcgccat cgaggcccag cagcacctgc tgcagctgac cgtgtggggc 1680
atcaagcagc tgcaggcccg cgtgctggcc gtggagcgct acctgaagga ccagcagctg 1740
ctgggcatct ggggctgcag cggcaagctg atctgcacca ccgccgtgcc ctggaacgcc 1800
agctggagca acaagagcct ggaccagatc tggaacaaca tgacctggat ggagtgggag 1860
cgcgagatcg acaactacac caacctgatc tacaccctga tcgaggagag ccagaaccag 1920
caggagaaga acgagcagga gctgctggag ctggacaagt gggccagcct gtggaactgg 1980
ttegacatea geaagtgget gtggtacate aagatettea teatgategt gggeggeetg 2040
gtgggcctgc gcatcgtgtt caccgtgctg agcatcgtga accgcgtgcg ccagggctac 2100
agccccctga gcttccagac ccgcttcccc gcccccgcg gccccgaccg ccccgagggc 2160
atcgaggagg agggcggcga gcgcgaccgc gaccgcagca gccccctggt gcacggcctg 2220
ctggccctga tctgggacga cctgcgcagc ctgtgcctgt tcagctacca ccgcctgcgc 2280
gacctgatec tgategeege cegeategtg gagetgetgg geegeegegg etgggaggee 2340
ctgaagtact ggggcaacct gctgcagtac tggatccagg agctgaagaa cagcgccgtg 2400
agcotyttcg acqccatcgc catcgccgtg gccgagggca ccgaccgcat catcgaggtg 2460
geccagegea teggeegege etteetgeae atecceegee geateegeea gggettegag 2520
cgcgccctgc tgtaactcga g
<210> 10
<211> 2541
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: Arg426-Gly431
<400> 10
gaattcgcca ccatggatgc aatgaagaga gggctctgct gtgtgctgct gctgtgtgga 60
```

gcagtcttcg tttcgcccag cgccgtggag aagctgtggg tgaccgtgta ctacggcgtg 120 cccgtgtgga aggaggccac caccaccctg ttctgcgcca gcgacgccaa ggcctacgac 180

```
accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 240
gagatcgtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtggag 300
cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaagctg 360
accccctgt gcgtgaccct gcactgcacc aacctgaaga acgccaccaa caccaagagc 420
agcaactgga aggagatgga ccgcggcgag atcaagaact gcagcttcaa ggtgaccacc 480
agcatcegea acaagatgea gaaggagtac geeetgttet acaagetgga egtggtgeee 540
atcgacaacg acaacaccag ctacaagctg atcaactgca acaccagcgt gatcacccag 600
gcctgcccca aggtgagctt cgagcccatc cccatccact actgcgcccc cgccggcttc 660
gccatcctga agtgcaacga caagaagttc aacggcagcg gcccctgcac caacgtgagc 720
acceptgeagt geacceaegg cateegeece gtggtgagea ceeagetget getgaaegge 780
agectggeeg aggagggegt ggtgateege agegagaact teacegacaa egecaagace 840
atcatcgtgc agctgaagga gagcgtggag atcaactgca cccgccccaa caacaacacc 900
cgcaagagca tcaccatcgg ccccggccgc gccttctacg ccaccggcga catcatcggc 960
gacatcogcc aggcccactg caacatcagc ggcgagaagt ggaacaacac cctgaagcag 1020
atogtgacca agotgoaggo coagttoggo aacaagacca togtgttoaa goagagcago 1080
ggcggcgacc ccgagatcgt gatgcacagc ttcaactgcg gcggcgagtt cttctactgc 1140
aacagcaccc agctgttcaa cagcacctgg aacaacacca teggecccaa caacaccaac 1200
ggcaccatca ccctgccctg ccgcatcaag cagatcatca accgcggcgg cggcaaggcc 1260
atgtacgccc cocccatccg cggccagate cgctgcagca gcaacatcac cggcctgctg 1320
ctgaccegeg acggcggcaa ggagatcagc aacaccaccg agatcttccg ccccggcggc 1380
ggcgacatgc gcgacaactg gcgcagcgag ctgtacaagt acaaggtggt gaagatcgag 1440
cccctgggcg tggccccac caaggccaag cgccgcgtgg tgcagcgcga gaagcgcgcc 1500
gtgaccetgg gegecatgtt cetgggette etgggegeeg eeggeageac catgggegee 1560
cgcagcctga ccctgaccgt gcaggcccgc cagctgctga gcggcatcgt gcagcagcag 1620
aacaacctgc tgcgcgccat cgaggcccag cagcacctgc tgcagctgac cgtgtggggc 1680
atcaagcagc tgcaggcccg cgtgctggcc gtggagcgct acctgaagga ccagcagctg 1740
ctgggcatct ggggctgcag cggcaagctg atctgcacca ccgccgtgcc ctggaacgcc 1800
agctggagca acaagagcct ggaccagatc tggaacaaca tgacctggat ggagtgggag 1860
cgcgagatcg acaactacac caacctgatc tacaccetga tcgaggagag ccagaaccag 1920
caggagaaga acgagcagga gctgctggag ctggacaagt gggccagcct gtggaactgg 1980
ttcgacatca gcaagtggct gtggtacatc aagatcttca tcatgatcgt gggcggcctg 2040
gtgggcctgc gcatcgtgtt caccgtgctg agcatcgtga accgcgtgcg ccagggctac 2100
agccccctga gcttccagac ccgcttcccc gcccccgcg gccccgaccg ccccgagggc 2160
atcgaggagg agggcggcga gcgcgaccgc gaccgcagca gccccctggt gcacggcctg 2220
ctggccctga tctgggacga cctgcgcagc ctgtgcctgt tcagctacca ccgcctgcgc 2280
gacctgatcc tgatcgccgc ccgcatcgtg gagctgctgg gccgccgcgg ctgggaggcc 2340
ctgaagtact ggggcaacct gctgcagtac tggatccagg agctgaagaa cagcgccgtg 2400
agectgttcg acgccatcgc catcgccgtg gccgagggca ccgaccgcat catcgaggtg 2460
gcccagegca teggeegege etteetgeae ateceeegee gcateegeea gggettegag 2520
cgcgccctgc tgtaactcga g
<210> 11
<211> 2541
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: Arg426-Gly431B
<400> 11
gaattegeca ecatggatge aatgaagaga gggetetget gtgtgetget getgtgtgga 60
gcagtcttcg tttcgcccag cgccgtggag aagctgtggg tgaccgtgta ctacggcgtg 120
cccgtgtgga aggaggccac caccacctg ttctgcgcca gcgacgccaa ggcctacgac 180
accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 240
gagatcgtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtggag 300
cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaagctg 360
accccctgt gcgtgaccct gcactgcacc aacctgaaga acgccaccaa caccaagagc 420
agcaactgga aggagatgga ccgcggcgag atcaagaact gcagcttcaa ggtgaccacc 480
```

agcatcogca acaagatgca gaaggagtac gccctgttct acaagctgga cgtggtgccc 540

```
ategacaacg acaacaccag ctacaagctg atcaactgca acaccagcgt gatcacccag 600
geetgeecca aggtgagett egageecate eccatecaet actgegeece egeeggette 660
gccatcctga agtgcaacga caagaagttc aacggcagcg gcccctgcac caacgtgagc 720
acceptgcagt gcacccacgg catccgcccc gtggtgagca cccagctgct gctgaacggc 780
agcctggccg aggaggcgt ggtgatccgc agcgagaact tcaccgacaa cgccaagacc 840
atcatcgtgc agctgaagga gagcgtggag atcaactgca cccgccccaa caacaacacc 900
cgcaagagca tcaccategg ccceggeege gccttctaeg ccaceggega catcategge 960
gacatccgcc aggcccactg caacatcagc ggcgagaagt ggaacaacac cctgaagcag 1020
atogtgacca agotgcaggo coagttoggo aacaagacca togtgttcaa goagagcago 1080
ggcggcgacc ccgagatcgt gatgcacagc ttcaactgcg gcggcgagtt cttctactgc 1140
aacagcaccc agctgttcaa cagcacctgg aacaacacca tcggccccaa caacaccaac 1200
ggcaccatca ccctgccctg ccgcatcaag cagatcatca accgcggcag cggcaaggcc 1260
atgtacgecc eccecatecg eggecagate egetgeagea geaacateac eggeetgetg 1320
ctgaccegcg acggcggcaa ggagatcagc aacaccaccg agatcttccg ccccggcggc 1380
ggcgacatgc gcgacaactg gcgcagcgag ctgtacaagt acaaggtggt gaagatcgag 1440
cccctgggcg tggcccccac caaggccaag cgccgcgtgg tgcagcgcga gaagcgcgcc 1500
gtgaccetgg gegecatgtt cetgggette etgggegeeg eeggeageac eatgggegee 1560
cgcagcctga ccctgaccgt gcaggcccgc cagctgctga gcggcatcgt gcagcagcag 1620
aacaacctgc tgcgcgccat cgaggcccag cagcacctgc tgcagctgac cgtgtggggc 1680
atcaagcagc tgcaggcccg cgtgctggcc gtggagcgct acctgaagga ccagcagctg 1740
ctgggcatct ggggctgcag cggcaagctg atctgcacca ccgccgtgcc ctggaacgcc 1800
agctggagca acaagagcct ggaccagatc tggaacaaca tgacctggat ggagtgggag 1860
cgcgagatcg acaactacac caacctgatc tacaccctga tcgaggagag ccagaaccag 1920
caggagaaga acgagcagga gctgctggag ctggacaagt gggccagcct gtggaactgg 1980
ttcgacatca gcaagtggct gtggtacatc aagatcttca tcatgatcgt gggcggcctg 2040
gtgggcctgc gcatcgtgtt caccgtgctg agcatcgtga accgcgtgcg ccagggctac 2100
ageccettga gettecagae cegetteece gecceegeg geecegaeeg eccegaggge 2160
ategaggagg agggeggega gegegacege gacegeagea geeceetggt geaeggeetg 2220
ctggccctga tctgggacga cctgcgcagc ctgtgcctgt tcagctacca ccgcctgcgc 2280
gacetgatee tgategeege cegeategtg gagetgetgg geegeegegg etgggaggee 2340
ctgaagtact ggggcaacct gctgcagtac tggatccagg agctgaagaa cagcgccgtg 2400
agcotgttcg acgccatcgc catcgccgtg gccgagggca ccgaccgcat catcgaggtg 2460
geccagegea teggeegege etteetgeae atecceegee geateegeea gggettegag 2520
cgcgccctgc tgtaactcga g
<210> 12
<211> 2541
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: Arg426-Lys432
<400> 12
gaattegeea eeatggatge aatgaagaga gggetetget gtgtgetget getgtgtgga 60
gcagtcttcg tttcgcccag cgccgtggag aagctgtggg tgaccgtgta ctacggcgtg 120
cccgtgtgga aggaggccac caccacctg ttctgcgcca gcgacgccaa ggcctacgac 180
accgaggtge acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 240
gagategtge tggagaaegt gacegagaae tteaacatgt ggaagaacaa catggtggag 300
cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaagctg 360
acceccetgt gegtgaccet geactgeace aacetgaaga acgeeaceaa caccaagage 420
agcaactgga aggagatgga ccgcggcgag atcaagaact gcagcttcaa ggtgaccacc 480
agcatccgca acaagatgca gaaggagtac gccctgttct acaagctgga cgtggtgccc 540
ategacaacg acaacaccag ctacaagetg ateaactgca acaccagegt gateacccag 600
geetgeeeca aggtgagett egageeeate eccateeact actgegeeec egeeggette 660
gccatcctga agtgcaacga caagaagttc aacggcagcg gcccctgcac caacgtgagc 720
acceptgcagt gcacccacgg catccgcccc gtggtgagca cccagctgct gctgaacggc 780
agcctggccg aggagggcgt ggtgatccgc agcgagaact tcaccgacaa cgccaagacc 840
atcatcgtgc agctgaagga gagcgtggag atcaactgca cccgccccaa caacaacacc 900
```

```
cgcaagagca tcaccatcgg ccccggccgc gccttctacg ccaccggcga catcatcggc 960
gacatccgcc aggcccactg caacatcagc ggcgagaagt ggaacaacac cctgaagcag 1020
atogtgacca agotgcaggo coagttoggo aacaagacca togtgttoaa gcagagcago 1080
ggcggcgacc ccgagatcgt gatgcacagc ttcaactgcg gcggcgagtt cttctactgc 1140
aacagcaccc agctgttcaa cagcacctgg aacaacacca tcggccccaa caacaccaac 1200
ggcaccatca ccctgccctg ccgcatcaag cagatcatca accgcggcgg caacaaggcc 1260
atgtacgccc cccccatccg cggccagatc cgctgcagca gcaacatcac cggcctgctg 1320
ctgacccgcg acggcggcaa ggagatcagc aacaccaccg agatcttccg ccccggcggc 1380
ggcgacatgc gcgacaactg gcgcagcgag ctgtacaagt acaaggtggt gaagatcgag 1440
cccctgggeg tggccccac caaggccaag cgccgcgtgg tgcagcgcga gaagcgcgcc 1500
gtgaccetgg gegecatgtt cetgggette etgggegeeg ceggeageac catgggegee 1560
cgcagcctga ccctgaccgt gcaggcccgc cagctgctga gcggcatcgt gcagcagcag 1620
aacaacctgc tgcgcgccat cgaggcccag cagcacctgc tgcagctgac cgtgtggggc 1680
atcaagcagc tgcaggcccg cgtgctggcc gtggagcgct acctgaagga ccagcagctg 1740
ctgggcatct ggggctgcag cggcaagctg atctgcacca ccgccgtgcc ctggaacgcc 1800
agctggagca acaagagcct ggaccagatc tggaacaaca tgacctggat ggagtgggag 1860
cgcgagatcg acaactacac caacctgatc tacaccctga tcgaggagag ccagaaccag 1920
caggagaaga acgagcagga gctgctggag ctggacaagt gggccagcct gtggaactgg 1980
ttcgacatca gcaagtggct gtggtacatc aagatcttca tcatgatcgt gggcggcctg 2040
gtgggcctgc gcatcgtgtt caccgtgctg agcatcgtga accgcgtgcg ccagggctac 2100
agreecetga getteeagae cegetteece geceeegeg geceegaeeg eecegaggge 2160
atcgaggagg agggcggcga gcgcgaccgc gaccgcagca gccccctggt gcacggcctg 2220
ctggccctga tctgggacga cctgcgcagc ctgtgcctgt tcagctacca ccgcctgcgc 2280
gacctgatec tgategeege cegeategtg gagetgetgg geegeegegg etggggggee 2340
ctgaagtact ggggcaacct gctgcagtac tggatccagg agctgaagaa cagcgccgtg 2400
agcotytteg acgocatege categoegtg geogagggea cegacegeat categaggtg 2460
geccagegea teggeegege etteetgeae atecceegee geateegeea gggettegag 2520
cgcgccctgc tgtaactcga g
<210> 13
<211> 2535
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: Asn425-Lys432
<400> 13
gaattegeca ceatggatge aatgaagaga gggetetget gtgtgetget getgtgtgga 60
gcagtcttcg tttcgcccag cgccgtggag aagctgtggg tgaccgtgta ctacggcgtg 120
cccgtgtgga aggaggccac caccaccctg ttctgcgcca gcgacgccaa ggcctacgac 180
accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 240
gagategtge tggagaaegt gaeegagaae tteaacatgt ggaagaaeaa catggtggag 300
cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaagctg 360
acceccetgt gegtgaceet geactgeace aacetgaaga aegecaceaa caccaagage 420
agcaactgga aggagatgga ccgcggcgag atcaagaact gcagcttcaa ggtgaccacc 480
ageateegea acaagatgea gaaggagtae geeetgttet acaagetgga egtggtgeee 540
atcgacaacg acaacaccag ctacaagctg atcaactgca acaccagcgt gatcacccag 600
gcctgcccca aggtgagctt cgagcccatc cccatccact actgcgcccc cgccggcttc 660
gccatcctga agtgcaacga caagaagttc aacggcagcg gcccctgcac caacgtgagc 720
accetteagt geacceaces cateegeeee gtesteagea cecaseteet getsaacese 780
agcotggccg aggagggcgt ggtgatccgc agcgagaact tcaccgacaa cgccaagacc 840
atcategtge agetgaagga gagegtggag atcaaetgea eeegeeecaa caacaacaee 900
cgcaagagca tcaccatcgg ccccggccgc gccttctacg ccaccggcga catcatcggc 960
gacateegee aggeeeactg caacateage ggegagaagt ggaacaacae eetgaageag 1020
atcgtgacca agctgcaggc ccagttcggc aacaagacca tcgtgttcaa gcagagcagc 1080
ggcggcgacc ccgagatcgt gatgcacagc ttcaactgcg gcggcgagtt cttctactgc 1140
aacagcaccc agctgttcaa cagcacctgg aacaacacca tcggccccaa caaccaac 1200
```

ggcaccatca ccctgccctg ccgcatcaag cagatcatca acgcccccaa ggccatgtac 1260

```
gecececca teegeggeea gateegetge ageageaaca teaceggeet getgetgaee 1320
cgcgacggcg gcaaggagat cagcaacacc accgagatct tccgccccgg cggcggcgac 1380
atgcgcgaca actggcgcag cgagctgtac aagtacaagg tggtgaagat cgagccctg 1440
ggcgtggccc ccaccaaggc caagcgccgc gtggtgcagc gcgagaagcg cgccgtgacc 1500
ctgggcgcca tgttcctggg cttcctgggc gccgccggca gcaccatggg cgcccgcagc 1560
ctgaccctga ccgtgcaggc ccgccagctg ctgagcggca tcgtgcagca gcagaacaac 1620
ctgctgcgcg ccatcgaggc ccagcagcac ctgctgcagc tgaccgtgtg gggcatcaag 1680
cagetgeagg ecegegtget ggeegtggag egetacetga aggaceagea getgetggge 1740
atotggggct gcagcggcaa gctgatctgc accaccgccg tgccctggaa cgccagctgg 1800
agcaacaaga gcctggacca gatctggaac aacatgacct ggatggagtg ggagcgcgag 1860
atcgacaact acaccaacct gatctacacc ctgatcgagg agagccagaa ccagcaggag 1920
aagaacgagc aggagctgct ggagctggac aagtgggcca gcctgtggaa ctggttcgac 1980
atcagcaagt ggctgtggta catcaagatc ttcatcatga tcgtgggcgg cctggtgggc 2040
ctgcgcatcg tgttcaccgt gctgagcatc gtgaaccgcg tgcgccaggg ctacagcccc 2100
ctgagettee agaceegett eccegeecee egeggeeceg acegeecega gggcategag 2160
gaggaggeg gegagegega cegegaeege ageageeece tggtgeaegg cetgetggee 2220
ctgatctggg acgacctgcg cagcctgtgc ctgttcagct accaccgcct gcgcgacctg 2280
atcctqatcq ccqcccqcat cgtggagctg ctgggccgcc gcggctggga ggccctgaag 2340
tactggggca acctgctgca gtactggatc caggagctga agaacagcgc cgtgagcctg 2400
ttcgacgcca tcgccatcgc cgtggccgag ggcaccgacc gcatcatcga ggtggcccag 2460
cgcatcggcc gcgccttcct gcacatcccc cgccgcatcc gccagggctt cgagcgcgcc 2520
ctgctgtaac tcgag
<210> 14
<211> 2529
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: Ile424-Ala433
<400> 14
gaattcgcca ccatggatgc aatgaagaga gggctctgct gtgtgctgct gctgtgtgga 60
gcaqtetteq ttteqeecag egceqtggag aagetgtggg tgacegtgta etaeggegtg 120
cccgtgtgga aggaggccac caccacctg ttctgcgcca gcgacgccaa ggcctacgac 180
accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 240
gagategtge tggagaaegt gacegagaae tteaacatgt ggaagaacaa catggtggag 300
cagatgeacg aggacateat cageetgtgg gaccagagee tgaageeetg egtgaagetg 360
accecetgt gegtgaceet geactgeace aacetgaaga acgeeaceaa caccaagage 420
agcaactgga aggagatgga ccgcggcgag atcaagaact gcagcttcaa ggtgaccacc 480
agcateegea acaagatgea gaaggagtae geeetgttet acaagetgga egtggtgeee 540
ategacaacg acaacaccag ctacaagetg ateaactgca acaccagegt gateacccag 600
geetgeecca aggtgagett egageecate eccatecaet actgegeece egeeggette 660
gccatcctga agtgcaacga caagaagttc aacggcagcg gcccctgcac caacgtgagc 720
acceptgeagt geacceaegg cateegeece gtggtgagea ceeagetget getgaaegge 780
agectggeeg aggagggegt ggtgateege agegagaaet teacegacaa egecaagaee 840
atcatcgtgc agctgaagga gagcgtggag atcaactgca cccgccccaa caacaacacc 900
cgcaagagca tcaccatcgg ccccggccgc gccttctacg ccaccggcga catcatcggc 960
gacatccgcc aggcccactg caacatcagc ggcgagaagt ggaacaacac cctgaagcag 1020
atcqtgacca agctgcaggc ccagttcggc aacaagacca tcgtgttcaa gcagagcagc 1080
ggcggcgacc ccgagatcgt gatgcacagc ttcaactgcg gcggcgagtt cttctactgc 1140
aacagcaccc agctgttcaa cagcacctgg aacaacacca tcggccccaa caacaccaac 1200
ggcaccatca ccctgccctg ccgcatcaag cagatcatcg gcggcgccat gtacgccccc 1260
cccatccgcg gccagatccg ctgcagcagc aacatcaccg gcctgctgct gacccgcgac 1320
ggcggcaagg agatcagcaa caccaccgag atcttccgcc ccggcggcgg cgacatgcgc 1380
gacaactggc gcagcgagct gtacaagtac aaggtggtga agatcgagcc cctgggcgtg 1440
gcccccacca aggccaagcg ccgcgtggtg cagcgcgaga agcgcgccgt gaccctgggc 1500
gccatgttcc tgggcttcct gggcgccgcc ggcagcacca tgggcgcccg cagcctgacc 1560
```

ctgaccgtgc aggcccgcca gctgctgagc ggcatcgtgc agcagcagaa caacctgctg 1620

```
cgcgccatcg aggcccagca gcacctgctg cagctgaccg tgtggggcat caagcagctg 1680
caggcccgcg tgctggccgt ggagcgctac ctgaaggacc agcagctgct gggcatctgg 1740
ggctgcagcg gcaagctgat ctgcaccacc gccgtgccct ggaacgccag ctggagcaac 1800
aagageetgg accagatetg gaacaacatg acctggatgg agtgggageg egagategae 1860
aactacacca acctgatcta caccctgatc gaggagagcc agaaccagca ggagaagaac 1920
gagcaggagc tgctggagct ggacaagtgg gccagcctgt ggaactggtt cgacatcagc 1980
aagtggctgt ggtacatcaa gatcttcatc atgatcgtgg gcggcctggt gggcctgcgc 2040
atogtgttca cogtgctgag catogtgaac cgcgtgcgcc agggctacag ccccctgagc 2100
ttccagaccc gcttccccgc ccccgcggc cccgaccgcc ccgagggcat cgaggaggag 2160
ggcggcgagc gcgaccgcga ccgcagcagc cccctggtgc acggcctgct ggccctgatc 2220
tgggacgacc tgcgcagcct gtgcctgttc agctaccacc gcctgcgcga cctgatcctg 2280
ategeegeee geategtgga getgetggge egeegegget gggaggeeet gaagtactgg 2340
ggcaacctgc tgcagtactg gatccaggag ctgaagaaca gcgccgtgag cctgttcgac 2400
gccatcgcca tcgccgtggc cgagggcacc gaccgcatca tcgaggtggc ccagcgcatc 2460
ggccgcgcct tcctgcacat cccccgccgc atccgccagg gcttcgagcg cgccctgctg 2520
taactcgag
<210> 15
<211> 2523
<212> DNA
<213> Artificial Sequence
<223> Description of Artificial Sequence: Ile423-Met434
<400> 15
gaattegeca ceatggatge aatgaagaga gggetetget gtgtgetget getgtgtgga 60
gcagtetteg tttegeccag egeegtggag aagetgtggg tgaeegtgta etaeggegtg 120
cccgtgtgga aggaggccac caccacctg ttctgcgcca gcgacgccaa ggcctacgac 180
accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 240
gagatcgtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtggag 300
cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaagctg 360
acceccetgt gegtgaccet geactgeace aacetgaaga aegecaceaa caccaagage 420
agcaactgga aggagatgga ccgcggcgag atcaagaact gcagcttcaa ggtgaccacc 480
agcatccgca acaagatgca gaaggagtac gccctgttct acaagctgga cgtggtgccc 540
atcgacaacg acaacaccag ctacaagctg atcaactgca acaccagcgt gatcacccag 600
gcctgcccca aggtgagctt cgagcccatc cccatccact actgcgcccc cgccggcttc 660
gccatcctga agtgcaacga caagaagttc aacggcagcg gcccctgcac caacgtgagc 720
acceptgeagt geacceaegg cateegeece gtggtgagea ceeagetget getgaaegge 780
agcetggccg aggagggcgt ggtgatccgc agcgagaact tcaccgacaa cgccaagacc 840
atcatcgtgc agctgaagga gagcgtggag atcaactgca cccgccccaa caacaacacc 900
cgcaagagca tcaccatcgg ccccggcegc gccttctacg ccaccggega catcatcggc 960
gacatccgcc aggcccactg caacatcagc ggcgagaagt ggaacaacac cctgaagcag 1020
atogtgacca agetgeagge ceagttegge aacaagacca tegtgtteaa geagageage 1080
ggcggcgacc ccgagatcgt gatgcacagc ttcaactgcg gcggcgagtt cttctactgc 1140
aacagcaccc agctgttcaa cagcacctgg aacaacacca tcggccccaa caacaccaac 1200
ggcaccatca ccctgccctg ccgcatcaag cagatcggcg gcatgtacgc cccccccatc 1260
cgcggccaga tccgctgcag cagcaacatc accggcctgc tgctgacccg cgacggcggc 1320
aaggagatca gcaacaccac cgagatcttc cgccccggcg gcggcgacat gcgcgacaac 1380
tggcgcagcg agctgtacaa gtacaaggtg gtgaagatcg agcccctggg cgtggccccc 1440
accaaggcca agegeegegt ggtgcagege gagaagegeg eegtgaceet gggegeeatg 1500
ttcctgggct tcctgggcgc cgccggcagc accatgggcg cccgcagcct gaccctgacc 1560
gtgcaggccc gccagctgct gagcggcatc gtgcagcagc agaacaacct gctgcgcgcc 1620
ategaggeec ageageacet getgeagetg acegtgtggg geateaagea getgeaggee 1680
cgcgtgctgg ccgtggagcg ctacctgaag gaccagcagc tgctgggcat ctggggctgc 1740
ageggeaage tgatetgeae cacegeegtg ceetggaaeg ceagetggag caacaagage 1800
ctggaccaga tctggaacaa catgacctgg atggagtggg agcgcgagat cgacaactac 1860
accaacctga tctacaccct gatcgaggag agccagaacc agcaggagaa gaacgagcag 1920
gagctgctgg agctggacaa gtgggccagc ctgtggaact ggttcgacat cagcaagtgg 1980
```

```
ctgtggtaca tcaagatctt catcatgatc gtgggcggcc tggtgggcct gcgcatcgtg 2040
ttcaccgtgc tgagcatcgt gaaccgcgtg cgccagggct acagccccct gagcttccag 2100
accegettee eegeceeeeg eggeeeegae egeceegagg geategagga ggagggegge 2160
gagegegace gegacegeag cagececetg gtgcaeggee tgctggccct gatetgggac 2220
gacctgegea gcctgtgect gttcagctac caccgcctgc gcgacctgat cctgatcgcc 2280
georgeateg tggagetget gggeegeege ggetgggagg ceetgaagta etggggeaac 2340
ctgctgcagt actggatcca ggagctgaag aacagcgccg tgagcctgtt cgacgccatc 2400
gccatcgccg tggccgaggg caccgaccgc atcatcgagg tggcccagcg catcggccgc 2460
geetteetge acateceeg eegeateege cagggetteg agegegeect getgtaacte 2520
gag
<210> 16
<211> 2517
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: Gln422-Tyr435
gaattcgcca ccatggatgc aatgaagaga gggctctgct gtgtgctgct gctgtgtgga 60
gcagtcttcg tttcgcccag cgccgtggag aagctgtggg tgaccgtgta ctacggcgtg 120
cccgtgtgga aggaggccac caccaccctg ttctgcgcca gcgacgccaa ggcctacgac 180
accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 240
gagatcgtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtggag 300
cagatgcacg aggacateat cageetgtgg gaccagagee tgaageeetg egtgaagetg 360
accccctgt gcgtgaccct gcactgcacc aacctgaaga acgccaccaa caccaagagc 420
agcaactgga aggagatgga ccgcggcgag atcaagaact gcagcttcaa ggtgaccacc 480
agcatccgca acaagatgca gaaggagtac gccctgttct acaagctgga cgtggtgccc 540
ategacaacg acaacaccag ctacaagetg ateaactgca acaccagegt gateacccag 600
gcctgcccca aggtgagctt cgagcccatc cccatccact actgcgcccc cgccggcttc 660
gccatcctga agtgcaacga caagaagttc aacggcagcg gcccctgcac caacgtgagc 720
acceptqcaqt gcacccacgg catccgcccc gtggtgagca cccagctgct gctgaacggc 780
agcetggeeg aggagggegt ggtgateege agegagaact teacegaeaa egeeaagaee 840
atcatcgtgc agctgaagga gagcgtggag atcaactgca cccgccccaa caacaacacc 900
cgcaagagca tcaccatcgg ccccggccgc gccttctacg ccaccggcga catcatcggc 960
gacatccgcc aggcccactg caacatcagc ggcgagaagt ggaacaacac cctgaagcag 1020
ategtgacca agetgeagge ceagttegge aacaagacca tegtgtteaa geagageage 1080
ggcggcgacc ccgagatcgt gatgcacagc ttcaactgcg gcggcgagtt cttctactgc 1140
aacagcaccc agctgttcaa cagcacctgg aacaacacca tcggccccaa caacaccaac 1200
ggcaccatca coctgccctg ccgcatcaag cagggcggct acgcccccc catccgcggc 1260
cagateeget geageageaa cateaeegge etgetgetga eeegegaegg eggeaaggag 1320
atcagcaaca ccaccgagat cttccgcccc ggcggcggcg acatgcgcga caactggcgc 1380
agegagetgt acaagtacaa ggtggtgaag ategageeee tgggegtgge ceccaccaag 1440
gccaagcgcc gcgtggtgca gcgcgagaag cgcgccgtga ccctgggcgc catgttcctg 1500
ggetteetgg gegeegegg cageaceatg ggegeeegea geetgaeeet gaeegtgeag 1560
geoegecage tgetgagegg categtgeag cageagaaca acetgetgeg egecategag 1620
gcccagcagc acctgctgca gctgaccgtg tggggcatca agcagctgca ggcccgcgtg 1680
ctggccgtgg agcgctacct gaaggaccag cagctgctgg gcatctgggg ctgcagcggc 1740
aagctgatct gcaccaccgc cgtgccctgg aacgccagct ggagcaacaa gagcctggac 1800
cagatctgga acaacatgac ctggatggag tgggagcgcg agatcgacaa ctacaccaac 1860
ctgatctaca ccctgatcga ggagagccag aaccagcagg agaagaacga gcaggagctg 1920
ctggagctgg acaagtgggc cagcctgtgg aactggttcg acatcagcaa gtggctgtgg 1980
tacatcaaga tetteateat gategtggge ggeetggtgg geetgegeat egtgtteace 2040
gtgctgagca tcgtgaaccg cgtgcgccag ggctacagcc ccctgagctt ccagacccgc 2100
ttccccgccc cccgcggccc cgaccgcccc gagggcatcg aggaggaggg cggcgagcgc 2160
gaccgcgacc gcagcagccc cctggtgcac ggcctgctgg ccctgatctg ggacgacctg 2220
cgcagcctgt gcctgttcag ctaccaccgc ctgcgcgacc tgatcctgat cgccgcccgc 2280
atogtggage tgctgggccg ccgcggctgg gaggccctga agtactgggg caacctgctg 2340
```

```
cagtactgga tccaggagct gaagaacagc gccgtgagcc tgttcgacgc catcgccatc 2400
gccgtggccg agggcaccga ccgcatcatc gaggtggccc agcgcatcgg ccgcgccttc 2460
ctgcacatcc cccgccgcat ccgccagggc ttcgagcgcg ccctgctgta actcgag
<210> 17
<211> 2517
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: Gln422-Tyr435B
<400> 17
gaattegeea ceatggatge aatgaagaga gggetetget gtgtgetget getgtgtgga 60
gcagtcttcg tttcgcccag cgccgtggag aagctgtggg tgaccgtgta ctacggcgtg 120
cccgtgtgga aggaggccac caccaccctg ttctgcgcca gcgacgccaa ggcctacgac 180
accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 240
gagatcgtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtggag 300
cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaagctg 360
accecettgt gegtgaceet geactgeace aacetgaaga acgecaceaa caccaagage 420
agcaactgga aggagatgga ccgcggcgag atcaagaact gcagcttcaa ggtgaccacc 480
agcatcegca acaagatgca gaaggagtac gccctgttct acaagctgga cgtggtgccc 540
ategacaacg acaacaccag ctacaagetg ateaactgca acaccagegt gateacccag 600
geetgeecca aggtgagett egageecate eccatecaet aetgegeece egeeggette 660
gccatcctga agtgcaacga caagaagttc aacggcagcg gcccctgcac caacgtgagc 720
acceptgcagt geacceaegg cateegeece gtggtgagca ceeagetget getgaaegge 780
agcotggccg aggagggcgt ggtgatccgc agcgagaact tcaccgacaa cgccaagacc 840
atcatcgtgc agctgaagga gagcgtggag atcaactgca cccgccccaa caacaacc 900
cgcaagagca tcaccatcgg ccccggccgc gccttctacg ccaccggcga catcatcggc 960
gacatccgcc aggcccactg caacatcagc ggcgagaagt ggaacaacac cctgaagcag 1020
ategtgacca agetgeagge ceagttegge aacaagacca tegtgtteaa geagageage 1080
ggcggcgacc ccgagatcgt gatgcacagc ttcaactgcg gcggcgagtt cttctactgc 1140
aacagcaccc agctgttcaa cagcacctgg aacaacacca tcggccccaa caacaccaac 1200
ggcaccatca ccctgccctg ccgcatcaag caggccccct acgcccccc catccgcggc 1260
cagateeget geageageaa cateaeegge etgetgetga eeegegaegg eggeaaggag 1320
atcagcaaca ccaccgagat cttccgcccc ggcggcggcg acatgcgcga caactggcgc 1380
agcgagctgt acaagtacaa ggtggtgaag atcgagcccc tgggcgtggc ccccaccaag 1440
gccaagegee gegtggtgea gegegagaag egegeegtga eeetgggege eatgtteetg 1500
ggcttcctgg gcgccgccgg cagcaccatg ggcgcccgca gcctgaccct gaccgtgcag 1560
gecegecage tgetgagegg categtgeag cageagaaca acetgetgeg egecategag 1620
geccageage acetgetgea getgacegtg tggggeatea ageagetgea ggecegegtg 1680
ctggccgtgg agcgctacct gaaggaccag cagctgctgg gcatctgggg ctgcagcggc 1740
aagetgatet geaccacege egtgeeetgg aacgeeaget ggageaacaa gageetggae 1800
cagatctgga acaacatgac ctggatggag tgggagcgcg agatcgacaa ctacaccaac 1860
ctgatctaca ccctgatcga ggagagccag aaccagcagg agaagaacga gcaggagctg 1920
ctggagctgg acaagtgggc cagcctgtgg aactggttcg acatcagcaa gtggctgtgg 1980
tacatcaaga tetteateat gategtggge ggeetggtgg geetgegeat egtgtteace 2040
gtgctgagca tcgtgaaccg cgtgcgccag ggctacagcc ccctgagctt ccagacccgc 2100
ttccccgccc cccgcggccc cgaccgcccc gagggcatcg aggaggaggg cggcgagcgc 2160
gaccgcgacc gcagcagccc cctggtgcac ggcctgctgg ccctgatctg ggacgacctg 2220
cgcagcetgt gcctgttcag ctaccaccgc ctgcgcgacc tgatcctgat cgccgcccgc 2280
atcgtggagc tgctgggccg ccgcggctgg gaggccctga agtactgggg caacctgctg 2340
cagtactgga tccaggagct gaagaacagc gccgtgagcc tgttcgacgc catcgccatc 2400
gccgtggccg agggcaccga ccgcatcatc gaggtggccc agcgcatcgg ccgcgccttc 2460
ctgcacatcc cccgccgcat ccgccagggc ttcgagcgcg ccctgctgta actcgag
<210> 18
<211> 2322
<212> DNA
```

<213> Artificial Sequence <220> <223> Description of Artificial Sequence: Leu122-Ser199; Arg426-Gly431 <400> 18 gaattegeea ceatggatge aatgaagaga gggetetget gtgtgetget getgtgtgga 60 gcagtetteg tttegeccag egeegtggag aagetgtggg tgacegtgta etaeggegtg 120 cccqtqtqqa aggaggccac caccacctg ttctgcgcca gcgacgccaa ggcctacgac 180 accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 240 gagatcgtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtggag 300 cagatgcacg aggacatcat cageetgtgg gaccagagee tgaageeetg egtgaagetg 360 ggcaacagcg tgatcaccca ggcctgcccc aaggtgagct tcgagcccat ccccatccac 420 tactqcqccc ccgccggctt cgccatcctg aagtgcaacg acaagaagtt caacggcagc 480 ggcccctgca ccaacgtgag caccgtgcag tgcacccacg gcatccgccc cgtggtgagc 540 acceagetge tgetgaaegg cageetggee gaggagggeg tggtgateeg cagegagaac 600 ttcaccgaca acgccaagac catcatcgtg cagctgaagg agagcgtgga gatcaactgc 660 accegecca acaacaca cegeaagage atcaceateg geeeggeeg egeettetac 720 gccaccggcg acatcatcgg cgacatccgc caggcccact gcaacatcag cggcgagaag 780 tggaacaaca ccctgaagca gatcgtgacc aagctgcagg cccagttcgg caacaagacc 840 ategtgttea ageagageag eggeggegae eeegagateg tgatgeacag etteaactge 900 ggcggcgagt tottotactg caacagcaco cagotgttca acagcacotg gaacaacaco 960 ateggececa acaacaccaa eggeaccate accetgecet geegeateaa geagateate 1020 aaccgcggcg geggcaaggc catgtacgcc ccccccatcc geggccagat ccgctgcagc 1080 aqcaacatca coggootgot gotgaccogo gacggoggoa aggagatcag caacaccacc 1140 gagatettee geeeeggegg eggegacatg egegacaact ggegeagega getgtacaag 1200 tacaaggtgg tgaagatcga gcccctgggc gtggccccca ccaaggccaa gcgccgcgtg 1260 gtgcagcgcg agaagcgcgc cgtgaccctg ggcgccatgt tcctgggctt cctgggcgcc 1320 geoggeagea ceatgggege cogeageetg accetgaceg tgeaggeeeg ceagetgetg 1380 ageggeateg tgeageagea gaacaacetg etgegegeea tegaggeeea geageacetg 1440 ctgcagctga ccgtgtgggg catcaagcag ctgcaggccc gcgtgctggc cgtggagcgc 1500 tacctgaagg accagcaget getgggcate tggggetgca geggcaaget gatetgcace 1560 accgccgtgc cctggaacgc cagctggagc aacaagagcc tggaccagat ctggaacaac 1620 atgacctgga tggagtggga gcgcgagatc gacaactaca ccaacctgat ctacaccctg 1680 atcgaggaga gccagaacca gcaggagaag aacgagcagg agctgctgga gctggacaag 1740 tgggccagcc tgtggaactg gttcgacatc agcaagtggc tgtggtacat caagatcttc 1800 atcatgatcg tgggcggcct ggtgggcctg cgcatcgtgt tcaccgtgct gagcatcgtg 1860 aaccgcgtgc gccagggcta cagcccctg agcttccaga cccgcttccc cgcccccgc 1920 ggccccgacc gccccgaggg catcgaggag gagggcggcg agcgcgaccg cgaccgcagc 1980 agccccctgg tgcacggcct gctggccctg atctgggacg acctgcgcag cctgtgcctg 2040 ttcagctacc accepctgcg cgacctgatc ctgatcgccg cccgcatcgt ggagctgctg 2100 ggccgccgcg gctgggaggc cctgaagtac tggggcaacc tgctgcagta ctggatccag 2160 gagetgaaga acagegeegt gageetgtte gaegeeateg ceategeegt ggeegaggge 2220 accgaccgca tcatcgaggt ggcccagcgc atcggccgcg cettectgca catcccccgc 2280 cgcatccgcc agggcttcga gcgcgccctg ctgtaactcg ag <210> 19 <211> 2322 <212> DNA <213> Artificial Sequence <220> <223> Description of Artificial Sequence: Leu122-Ser199; Arg426-Lys432 <400> 19 gaattegeea ceatggatge aatgaagaga gggetetget gtgtgetget getgtgtgga 60

gcagtetteg tttegeccag egcegtggag aagetgtggg tgacegtgta etaeggegtg 120

```
cccgtgtgga aggaggccac caccaccetg ttctgcgcca gcgacgccaa ggcctacgac 180
accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 240
gagatogtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtggag 300
cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaagctg 360
ggcaacageg tgatcaccca ggcctgcccc aaggtgaget tegageccat ceccatecae 420
tactgegeee eegeeggett egecateetg aagtgeaacg acaagaagtt caacggeage 480
ggcccctgca ccaacgtgag caccgtgcag tgcacccacg gcatccgccc cgtggtgagc 540
acccagetge tgctgaacgg cageetggee gaggagggeg tggtgateeg cagegagaac 600
ttcaccgaca acgccaagac catcatcgtg cagctgaagg agagcgtgga gatcaactgc 660
accegececa acaacaacae eegcaagage atcaceateg geeeeggeeg egeettetae 720
gccaccggcg acatcatcgg cgacatccgc caggcccact gcaacatcag cggcgagaag 780
tggaacaaca ccctgaagca gatcgtgacc aagctgcagg cccagttcgg caacaagacc 840
atcgtgttca agcagagcag cggcggcgac cccgagatcg tgatgcacag cttcaactgc 900
ggcggcgagt tcttctactg caacagcacc cagctgttca acagcacctg gaacaacacc 960
ateggeecca acaacaccaa eggeaccate accetgeect geegeatcaa geagateate 1020
aaccgcggcg gcaacaaggc catgtacgcc cccccatcc gcggccagat ccgctgcagc 1080
agcaacatca coggootgot gotgaccogo gacggoggoa aggagatcag caacaccacc 1140
gagatettee geeeeggegg eggegacatg egegacaaet ggegeagega getgtacaag 1200
tacaaggtgg tgaagatcga gcccctgggc gtggccccca ccaaggccaa gcgccgcgtg 1260
gtgcagcgcg agaagcgcgc cgtgaccctg ggcgccatgt tcctgggctt cctgggcgcc 1320
geoggeagea ceatgggege eegeageetg accetgaceg tgeaggeeeg ceagetgetg 1380
agcggcatcg tgcagcagca gaacaacctg ctgcgcgcca tcgaggccca gcagcacctg 1440
ctgcagctga ccgtgtgggg catcaagcag ctgcaggccc gcgtgctggc cgtggagcgc 1500
tacctgaagg accagcagct gctgggcatc tggggctgca gcggcaagct gatctgcacc 1560
accgccgtgc cctggaacgc cagctggagc aacaagagcc tggaccagat ctggaacaac 1620
atgacctgga tggagtggga gcgcgagatc gacaactaca ccaacctgat ctacaccctg 1680
atcgaggaga gccagaacca gcaggagaag aacgagcagg agctgctgga gctggacaag 1740
tgggccagcc tgtggaactg gttcgacatc agcaagtggc tgtggtacat caagatette 1800
atcatgateg tgggcggcct ggtgggcctg cgcatcgtgt tcaccgtgct gagcatcgtg 1860
aaccgcgtgc gccagggcta cagccccctg agcttccaga cccgcttccc cgcccccgc 1920
ggccccgacc gccccgaggg catcgaggag gagggcggcg agcgcgaccg cgaccgcagc 1980
agccccctgg tgcacggcct gctggccctg atctgggacg acctgcgcag cctgtgcctg 2040
ttcagctacc accgcctgcg cgacctgatc ctgatcgccg cccgcatcgt ggagctgctg 2100
ggccgccgcg gctgggaggc cctgaagtac tggggcaacc tgctgcagta ctggatccag 2160
gagetgaaga acagegeegt gageetgtte gaegeeateg ecategeegt ggeegaggge 2220
accgaccgca tcatcgaggt ggcccagcgc atcggccgcg ccttcctgca catcccccgc 2280
cgcatccgcc agggcttcga gcgcgccctg ctgtaactcg ag
<210> 20
<211> 2322
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: Leu122-Ser199;
      Trp427-Gly431
<400> 20
gaattcgcca ccatggatgc aatgaagaga gggctctgct gtgtgctgct gctgtgtgga 60
gcagtcttcg tttcgcccag cgccgtggag aagctgtggg tgaccgtgta ctacggcgtg 120
cccgtgtgga aggaggccac caccaccctg ttctgcgcca gcgacgccaa ggcctacgac 180
accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 240
gagatcgtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtggag 300
cagatgcacg aggacatcat cagcotgtgg gaccagagco tgaagcootg cgtgaagotg 360
ggcaacagcg tgatcaccca ggcctgcccc aaggtgagct tcgagcccat ccccatccac 420
tactgegeee eegeeggett egecateetg aagtgeaacg acaagaagtt caacggeage 480
ggcccctgca ccaacgtgag caccgtgcag tgcacccacg gcatccgccc cgtggtgagc 540
acccaqctqc tqctqaacgg cagcctggcc gaggagggcg tggtgatccg cagcgagaac 600
ttcaccgaca acgccaagac catcatcgtg cagctgaagg agagcgtgga gatcaactgc 660
```

```
accegecca acaacacac cegeaagage atcaceateg geeeeggeeg egeettetae 720
gccaccggcg acatcatcgg cgacatccgc caggcccact gcaacatcag cggcgagaag 780
tggaacaaca ccctgaagca gatcgtgacc aagctgcagg cccagttcgg caacaagacc 840
atcgtgttca agcagagcag cggcggcgac cccgagatcg tgatgcacag cttcaactgc 900
ggcggcgagt tcttctactg caacagcacc cagetgttca acagcacctg gaacaacacc 960
atcggcccca acaacaccaa cggcaccatc accctgccct gccgcatcaa gcagatcatc 1020
aaccgctggg gcggcaaggc catgtacgcc cccccatcc gcggccagat ccgctgcagc 1080
agcaacatca ccggcctgct gctgacccgc gacggcggca aggagatcag caacaccacc 1140
gagatettee geeceggegg eggegacatg egegacaact ggegeagega getgtacaag 1200
tacaaggtgg tgaagatcga gcccctgggc gtggccccca ccaaggccaa gcgccgcgtg 1260
gtgcagcgcg agaagcgcgc cgtgaccctg ggcgccatgt tcctgggctt cctgggcgcc 1320
gccggcagca ccatgggcgc ccgcagcctg accctgaccg tgcaggcccg ccagctgctg 1380
ageggeateg tgeageagea gaacaacetg etgegegeea tegaggeeea geageacetg 1440
ctgcagctga ccgtgtgggg catcaagcag ctgcaggccc gcgtgctggc cgtggagcgc 1500
tacctgaagg accagcaget getgggcate tggggctgca geggcaaget gatetgcace 1560
accgccgtgc cctggaacgc cagctggagc aacaagagcc tggaccagat ctggaacaac 1620
atgacctgga tggagtggga gcgcgagatc gacaactaca ccaacctgat ctacaccctg 1680
atcgaggaga gccagaacca gcaggagaag aacgagcagg agctgctgga gctggacaag 1740
tgggccagcc tgtggaactg gttcgacatc agcaagtggc tgtggtacat caagatcttc 1800
atcatgatcg tgggcggcct ggtgggcctg cgcatcgtgt tcaccgtgct gagcatcgtg 1860
aaccgcgtgc gccagggcta cagcccctg agcttccaga cccgcttccc cgcccccgc 1920
ggccccgacc gccccgaggg catcgaggag gagggcggcg agcgccaccg cgaccgcagc 1980
agccccctgg tgcacggcct gctggccctg atctgggacg acctgcgcag cctgtgcctg 2040
ttcagctacc accgcctgcg cgacctgatc ctgatcgccg cccgcatcgt ggagctgctg 2100
ggccgccgcg gctgggaggc cctgaagtac tggggcaacc tgctgcagta ctggatccag 2160
gagetgaaga acagegeegt gageetgtte gaegeeateg ceategeegt ggeegaggge 2220
accgaccgca tcatcgaggt ggcccagcgc atcggccgcg cettectgca catcccccgc 2280
cgcatccgcc agggcttcga gcgcgccctg ctgtaactcg ag
<210> 21
<211> 2310
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: Lys121-Val200;
      Asn425-Lys432
<400> 21
gaattegeea ceatggatge aatgaagaga gggetetget gtgtgetget getgtgtgga 60
gcagtetteg tttegcccag egcegtggag aagetgtggg tgacegtgta etacggegtg 120
cccgtgtgga aggaggccac caccacctg ttctgcgcca gcgacgccaa ggcctacgac 180
accgaggtge acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 240
gagatcgtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtggag 300
cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaaggcc 360
cccqtqatca cccaqqcctq ccccaaqqtq agcttcqaqc ccatccccat ccactactqc 420
gccccgccg gcttcgccat cctgaagtgc aacgacaaga agttcaacgg cagcgcccc 480
tgcaccaacg tgagcaccgt gcagtgcacc cacggcatcc gccccgtggt gagcacccag 540
ctgctgctga acggcagcct ggccgaggag ggcgtggtga tccgcagcga gaacttcacc 600
gacaacgcca agaccatcat cgtgcagctg aaggagagcg tggagatcaa ctgcacccgc 660
cccaacaaca acaccegcaa gagcatcacc ateggeceeg geogegeett ctaegecace 720
ggcgacatca tcggcgacat ccgccaggcc cactgcaaca tcagcggcga gaagtggaac 780
aacaccctga agcagatcgt gaccaagctg caggcccagt tcggcaacaa gaccatcgtg 840
ttcaagcaga gcagcggcgg cgaccccgag atcgtgatgc acagcttcaa ctgcggcggc 900
gagttettet actgeaacag cacceagetg tteaacagea eetggaacaa caccategge 960
eccaacaaca ccaacggcac catcaccetg ccetgcegca tcaagcagat catcaacgcc 1020
cccaaggcca tgtacgcccc ccccatccgc ggccagatcc gctgcagcag caacatcacc 1080
ggcctgctgc tgacccgcga cggcggcaag gagatcagca acaccaccga gatcttccgc 1140
eccggeggeg gegacatgeg egacaactgg egeagegage tgtacaagta caaggtggtg 1200
```

```
aagategage ceetgggegt ggeeceeace aaggeeaage geegegtggt geagegegag 1260
aagegegeeg tgaccetggg egecatgtte etgggettee tgggegeege eggeageace 1320
atgggcgccc gcagcctgac cctgaccgtg caggcccgcc agctgctgag cggcatcgtg 1380
cagcagcaga acaacctgct gcgcgccatc gaggcccagc agcacctgct gcagctgacc 1440
gtgtggggca tcaagcagct gcaggcccgc gtgctggccg tggagcgcta cctgaaggac 1500
cagcagetge tgggcatetg gggetgeage ggcaagetga tetgeaceae egeegtgeee 1560
tggaacgcca gctggagcaa caagagcctg gaccagatct ggaacaacat gacctggatg 1620
gagtgggagc gcgagatcga caactacacc aacctgatct acaccctgat cgaggagagc 1680
cagaaccagc aggagaagaa cgagcaggag ctgctggagc tggacaagtg ggccagcctg 1740
tggaactggt tcgacatcag caagtggctg tggtacatca agatcttcat catgatcgtg 1800
ggcggcctgg tgggcctgcg catcgtgttc accgtgctga gcatcgtgaa ccgcgtgcgc 1860
cagggetaca geceetgag ettecagace egetteeceg ecceegegg eccegacege 1920
cccgagggca tcgaggagga gggcggcgag cgcgaccgcg accgcagcag ccccctggtg 1980
cacggcetge tggccetgat etgggacgae etgegeagee tgtgcctgtt cagetaceae 2040
egectgegeg acctgateet gategeegee egeategtgg agetgetggg eegeegegge 2100
tgggaggecc tgaagtactg gggcaacctg ctgcagtact ggatccagga gctgaagaac 2160
agegeegtga geetgttega egecategee ategeegtgg eegagggeac egacegeate 2220
atcgaggtgg cccagcgcat cggccgcgcc ttcctgcaca tcccccgccg catccgccag 2280
ggcttcgagc gcgccctgct gtaactcgag
<210> 22
<211> 2298
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: Vall20-Ile201;
      Ile424-Ala433
<400> 22
gaattegeea ceatggatge aatgaagaga gggetetget gtgtgetget getgtgtgga 60
gcagtcttcg tttcgcccag cgccgtggag aagctgtggg tgaccgtgta ctacggcgtg 120
cccgtgtgga aggaggccac caccacctg ttctgcgcca gcgacgccaa ggcctacgac 180
accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 240
gagategtge tggagaaegt gaeegagaae tteaacatgt ggaagaaeaa catggtggag 300
cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgggcggc 360
ateacceagg cetgeceeaa ggtgagette gageceatee ceatecaeta etgegeeece 420
geeggetteg ceateetgaa gtgeaacgae aagaagttea aeggeagegg eeeetgeace 480
aacgtgagca ccgtgcagtg cacccacggc atccgccccg tggtgagcac ccagctgctg 540
ctgaacggca gcctggccga ggagggcgtg gtgatccgca gcgagaactt caccgacaac 600
gccaagacca tcatcgtgca gctgaaggag agcgtggaga tcaactgcac ccgccccaac 660
aacaacacc gcaagagcat caccatcggc cccggccgcg ccttctacgc caccggcgac 720
atcatcggcg acatccgcca ggcccactgc aacatcagcg gcgagaagtg gaacaacacc 780
ctgaagcaga tcgtgaccaa gctgcaggcc cagttcggca acaagaccat cgtgttcaag 840
cagagcagcg gcggcgaccc cgagatcgtg atgcacagct tcaactgcgg cggcgagttc 900
ttctactgca acagcaccca gctgttcaac agcacctgga acaacaccat cggccccaac 960
aacaccaacg gcaccatcac cctgccctgc cgcatcaagc agatcatcgg cggcgccatg 1020
tacgccccc ccatccgcgg ccagatccgc tgcagcagca acatcaccgg cctgctgctg 1080
accegegacg geggeaagga gateageaac accacegaga tetteegeec eggeggegge 1140
gacatgegeg acaactggeg cagegagetg tacaagtaca aggtggtgaa gategageee 1200
ctgggcgtgg cccccaccaa ggccaagcgc cgcgtggtgc agcgcgagaa gcgcgccgtg 1260
accetgggcg ccatgtteet gggetteetg ggcgccgccg gcagcaccat gggcgcccgc 1320
agcotgacco tgacogtgca ggoocgocag otgotgagog gcatogtgca gcagcagaac 1380
aacctgctgc gcgccatcga ggcccagcag cacctgctgc agctgaccgt gtggggcatc 1440
aagcagctgc aggcccgcgt gctggccgtg gagcgctacc tgaaggacca gcagctgctg 1500
ggcatctggg gctgcagcgg caagctgatc tgcaccaccg ccgtgccctg gaacgccagc 1560
tggagcaaca agagcctgga ccagatctgg aacaacatga cctggatgga gtgggagcgc 1620
gagatcgaca actacaccaa cctgatctac accctgatcg aggagagcca gaaccagcag 1680
```

gagaagaacg agcaggagct gctggagctg gacaagtggg ccagcctgtg gaactggttc 1740

```
gacatcagca agtggctgtg gtacatcaag atcttcatca tgatcgtggg cggcctggtg 1800
ggectgegea tegtgtteac egtgetgage ategtgaace gegtgegeea gggetacage 1860
cccctgagct tccagacccg cttccccgcc ccccgcggcc ccgaccgccc cgagggcatc 1920
gaggaggagg gcggcgagcg cgaccgcgac cgcagcagcc ccctggtgca cggcctgctg 1980
gccctgatct gggacgacct gcgcagcctg tgcctgttca gctaccaccg cctgcgcgac 2040
ctgatcctga tcgccgcccg catcgtggag ctgctgggcc gccgcggctg ggaggccctg 2100
aagtactggg gcaacctgct gcagtactgg atccaggagc tgaagaacag cgccgtgagc 2160
ctgttcgacg ccatcgccat cgccgtggcc gagggcaccg accgcatcat cgaggtggcc 2220
cagegeateg geogegeett cetgeacate eccegeegea teegeeaggg ettegagege 2280
gccctgctgt aactcgag
<210> 23
<211> 2298
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence:
      Val120-Ile201B; Ile424-Ala433
<400> 23
gaattegeea eeatggatge aatgaagaga gggetetget gtgtgetget getgtgtgga 60
geagtetteg tttegeceag egeegtggag aagetgtggg tgacegtgta etaeggegtg 120
cccgtgtgga aggaggccac caccaccetg ttetgcgcca gcgacgccaa ggcctacgac 180
accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 240
gagategtge tggagaaegt gaeegagaae tteaacatgt ggaagaaeaa catggtggag 300
cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgcccggc 360
atcacccagg cctgccccaa ggtgagcttc gagcccatcc ccatccacta ctgcgccccc 420
gccggcttcg ccatcctgaa gtgcaacgac aagaagttca acggcagcgg cccctgcacc 480
aacgtgagca ccgtgcagtg cacccacggc atccgccccg tggtgagcac ccagctgctg 540
ctgaacggca gcctggccga ggagggcgtg gtgatccgca gcgagaactt caccgacaac 600
gccaagacca tcatcgtgca gctgaaggag agcgtggaga tcaactgcac ccgccccaac 660
aacaacaccc gcaagagcat caccatcggc cccggccgcg ccttctacgc caccggcgac 720
atcatcggcg acatccgcca ggcccactgc aacatcagcg gcgagaagtg gaacaacacc 780
ctgaagcaga tcgtgaccaa gctgcaggcc cagttcggca acaagaccat cgtgttcaag 840
cagagcagcg geggegacce egagategtg atgeacaget teaactgegg eggegagtte 900
ttctactgca acagcaccca gctgttcaac agcacctgga acaacaccat cggccccaac 960
aacaccaacg gcaccatcac cctgccctgc cgcatcaagc agatcatcgg cggcgccatg 1020
tacgcccccc ccatccgcgg ccagatccgc tgcagcagca acatcaccgg cctgctgctg 1080
accegegacg geggeaagga gateagéaac accacegaga tetteegece eggeggegge 1140
gacatgcgcg acaactggcg cagcgagctg tacaagtaca aggtggtgaa gatcgagccc 1200
ctgggcgtgg cccccaccaa ggccaagcgc cgcgtggtgc agcgcgagaa gcgcgccgtg 1260
accetgggeg ceatgtteet gggetteetg ggegeegeeg geageaceat gggegeeege 1320
agcetgacce tgacegtgca ggeecgccag etgetgageg geategtgea geageagaae 1380
aacctgctgc gcgccatcga ggcccagcag cacctgctgc agctgaccgt gtggggcatc 1440
aagcagctgc aggcccgcgt gctggccgtg gagcgctacc tgaaggacca gcagctgctg 1500
ggcatctggg gctgcagcgg caagctgatc tgcaccaccg ccgtgccctg gaacgccagc 1560
tggagcaaca agagcctgga ccagatctgg aacaacatga cctggatgga gtgggagcgc 1620
gagategaca actacaceaa cetgatetae accetgateg aggagageea gaaceageag 1680
gagaagaacg agcaggagct gctggagctg gacaagtggg ccagcctgtg gaactggttc 1740
gacatcagca agtggctgtg gtacatcaag atcttcatca tgatcgtggg cggcctggtg 1800
ggcctgcgca tcgtgttcac cgtgctgagc atcgtgaacc gcgtgcgcca gggctacagc 1860
cccctgagct tecagacccg cttccccgcc ccccgcggcc ccgaccgccc cgagggcatc 1920
gaggaggagg gcggcgagcg cgaccgcgac cgcagcagcc ccctggtgca cggcctgctg 1980
gccctgatct gggacgacct gcgcagcctg tgcctgttca gctaccaccg cctgcgcgac 2040
ctgatcctga tcgccgcccg catcgtggag ctgctgggcc gccgcggctg ggaggccctg 2100
aagtactggg gcaacctgct gcagtactgg atccaggagc tgaagaacag cgccgtgagc 2160
ctgttcgacg ccatcgccat cgccgtggcc gagggcaccg accgcatcat cgaggtggcc 2220
```

cagegeateg geogegeett cetgeacate eccegeegea teegeeaggg ettegagege 2280

gcagtcttcg tttcgcccag cgccgtggag aagctgtggg tgaccgtgta ctacggcgtg 120 cccgtgtgga aggaggccac caccacctg ttctgcgcca gcgacgccaa ggcctacgac 180 accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccaq 240 gagatcgtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtggag 300 cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgggcggc 360 gccacccagg cctgccccaa ggtgagcttc gagcccatcc ccatccacta ctgcgccccc 420 gccggcttcg ccatcctgaa gtgcaacgac aagaagttca acggcagcgg cccctgcacc 480 aacgtgagca ccgtgcagtg cacccacggc atccgcccg tggtgagcac ccagctgctg 540 ctgaacggca gcctggccga ggagggcgtg gtgatccgca gcgagaactt caccgacaac 600 gccaagacca tcatcgtgca gctgaaggag agcgtggaga tcaactgcac ccgccccaac 660 aacaacaccc gcaagagcat caccateggc ceeggeegeg cettetaegc caceggegac 720 atcatcggcg acatccgcca ggcccactgc aacatcagcg gcgagaagtg gaacaacacc 780 ctgaagcaga tcgtgaccaa gctgcaggcc cagttcggca acaagaccat cgtgttcaag 840 cagageageg geggegaeec egagategtg atgeacaget teaactgegg eggegagtte 900 ttctactgca acagcaccca gctgttcaac agcacctgga acaacaccat cggccccaac 960 aacaccaacg gcaccatcac cctgccctgc cgcatcaagc agatcatcgg cggcgccatg 1020 tacgccccc ccatccgcgg ccagatccgc tgcagcagca acatcaccgg cctgctgctg 1080 accegegacg geggeaagga gateageaac accacegaga tetteegeec eggeggegge 1140 gacatgcgcg acaactggcg cagcgagctg tacaagtaca aggtggtgaa gatcgagccc 1200 ctgggcgtgg cccccaccaa ggccaagcgc cgcgtggtgc agcgcgagaa gcgcgccgtg 1260 accotgggcg coatgitect gggcttectg ggcgccgcg gcagcaccat gggcgcccgc 1320 agectgacce tgacegtgca ggecegecag etgetgageg geategtgca geageagaac 1380 aacctgctgc gcgccatcga ggcccagcag cacctgctgc agctgaccgt gtggggcatc 1440 aagcagctgc aggcccgcgt gctggccgtg gagcgctacc tgaaggacca gcagctgctq 1500 ggcatctggg gctgcagcgg caagctgatc tgcaccaccg ccgtgccctg gaacgccagc 1560 tggagcaaca agagcctgga ccagatctgg aacaacatga cctggatgga gtgggagcgc 1620 gagategaca actacaceaa cetgatetae accetgateg aggagageca qaaceaqeaq 1680 gagaagaacg agcaggagct gctggagctg gacaagtggg ccagcctgtg gaactggttc 1740 gacatcagca agtggctgtg gtacatcaag atcttcatca tgatcgtggg cggcctggtg 1800 ggcctgcgca tcgtgttcac cgtgctgagc atcgtgaacc gcgtgcgcca gggctacagc 1860 cocctgaget tecagacceg etteccegec coccgeggec cegacegece egagggeate 1920 gaggaggagg gcggcgagcg cgaccgcgac cgcagcagcc ccctggtgca cggcctgctq 1980 geeetgatet gggaegaeet gegeageetg tgeetgttea getaceaeeg eetgegegae 2040 ctgatcctga tcgccgcccg catcgtggag ctgctgggcc gccgcggctg ggaggccctg 2100 aagtactggg gcaacctgct gcagtactgg atccaggagc tgaagaacag cgccgtgagc 2160 ctgttcgacg ccatcgccat cgccgtggcc gagggcaccg accgcatcat cgaggtggcc 2220 cagcgcatcg geogegeett cetgcacate eccegeegea teegeeaggg ettegagege 2280 gccctgctgt aactcgag 2298

```
<210> 25
<211> 2358
<212> DNA
<213> Artificial Sequence
```

<223> Description of Artificial Sequence: Vall27-Asn195

```
<400> 25
 gaattcgcca ccatggatgc aatgaagaga gggctctgct gtgtgctgct gctgtgtgga 60
 gcagtcttcg tttcgcccag cgccgtggag aagctgtggg tgaccgtgta ctacggcgtg 120
 cccgtgtgga aggaggccac caccaccctg ttctgcgcca gcgacgccaa ggcctacgac 180
 accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 240
 gagatogtgo tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtggag 300
cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaagctg 360
accccctgt gcgtggggc agggaactgc aacaccagcg tgatcaccca ggcctgcccc 420
aaggtgaget tegageeeat ecceateeac tactgegeee eegeeggett egecateetg 480
aagtgcaacg acaagaagtt caacggcagc ggcccctgca ccaacgtgag caccgtgcag 540
tgcacccacg gcatccgccc cgtggtgagc acccagctgc tgctgaacgg cagcctggcc 600
gaggaggcg tggtgatccg cagcgagaac ttcaccgaca acgccaagac catcatcgtg 660
cagctgaagg agagcgtgga gatcaactgc accegececa acaacaacac cegcaagage 720
atcaccateg geoeggeeg egeettetae gecaceggeg acateategg egacateege 780
caggeceact geaacateag eggegagaag tggaacaaca eeetgaagea gategtgace 840
cccgagatcg tgatgcacag cttcaactgc ggcggcgagt tcttctactg caacagcacc 960
cagetgttea acageacetg gaacaacace ateggeecea acaacaceaa eggeaceate 1020
accetgeeet geegeateaa geagateate aacegetgge aggaggtggg caaggeeatg 1080
tacgccccc ccatccgcgg ccagatccgc tgcagcagca acatcaccgg cctgctgctg 1140
accegegacg geggeaagga gateageaac accacegaga tetteegeec eggeggegge 1200
gacatgogog acaactggog cagogagotg tacaagtaca aggtggtgaa gatogagooc 1260
ctgggcgtgg cccccaccaa ggccaagcgc cgcgtggtgc agcgcgagaa gcgcgccqtq 1320
accetgggeg ceatgtteet gggetteetg ggegeegeeg geageaceat gggegeeege 1380
agcetgacce tgacegtgca ggecegecag etgetgageg geategtgca geageagaac 1440
aacctgctgc gcgccatcga ggcccagcag cacctgctgc agctgaccgt gtggggcatc 1500
aagcagctgc aggcccgcgt gctggccgtg gagcgctacc tgaaggacca gcagctgctg 1560
ggcatctggg gctgcagcgg caagctgatc tgcaccaccg ccgtgccctg gaacgccagc 1620
tggagcaaca agagcctgga ccagatctgg aacaacatga cctggatgga gtgggagcgc 1680
gagategaca actacaceaa eetgatetae accetgateg aggagageea gaaceageag 1740
gagaagaacg agcaggagct gctggagctg gacaagtggg ccagcctgtg gaactggttc 1800
gacatcagca agtggctgtg gtacatcaag atcttcatca tgatcgtggg cggcctggtg 1860
ggcctgcgca tcgtgttcac cgtgctgagc atcgtgaacc gcgtgcgcca gggctacagc 1920
cccctgagct tccagacccg cttccccgcc ccccgcggcc ccgaccgccc cgagggcatc 1980
gaggaggagg gcggcgagcg cgaccgcgac cgcagcagcc ccctggtgca cggcctgctg 2040
geoetgatet gggacgacet gegeageetg tgeetgttea getaceaeeg eetgegegae 2100
ctgatcctga tcgccgcccg catcgtggag ctgctgggcc gccgcggctg ggaggccctg 2160
aagtactggg gcaacctgct gcagtactgg atccaggagc tgaagaacag cgccgtgagc 2220
ctgttcgacg ccatcgccat cgccgtggcc gagggcaccg accgcatcat cgagqtqqcc 2280
cagogcateg geogegeett cetgeacate eccegeegea teegeeaggg ettegagege 2340
gccctgctgt aactcgag
                                                                 2358
<210> 26
<211> 2352
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: Val127-Asn195;
      Arg426-Gly431
<400> 26
gaattcgcca ccatggatgc aatgaagaga gggctctgct gtgtgctgct gctgtgtgga 60
gcagtetteg tttegeccag egeegtggag aagetgtggg tgacegtgta etaeggegtg 120
cccgtgtgga aggaggccac caccaccctg ttctgcgcca gcgacgccaa ggcctacgac 180
accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 240
gagategtge tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtggag 300
cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaagctg 360
```

accecetgt gegtggggge agggaactge aacaceageg tgateaceca ggeetgeeec 420

aaggtgagct	tcgagcccat	ccccatccac	tactgcgccc	ccgccggctt	cgccatcctg	480
aagtgcaacg	acaagaagtt	caacggcagc	ggcccctgca	ccaacgtgag	caccgtgcag	540
tgcacccacg	gcatccgccc	cgtggtgagc	acccagetge	tgctgaacgg	cagcctggcc	600
	tggtgatccg					
cagctgaagg	agagcgtgga	gatcaactgc	acccgcccca	acaacaacac	ccgcaagagc	720
atcaccatcg	gccccggccg	cgccttctac	gccaccggcg	acatcatcgg	cgacatccgc	780
	gcaacatcag					840
	cccagttcgg					
	tgatgcacag					
	acagcacctg					
accctgccct	gccgcatcaa	gcagatcatc	aaccgcggcg	gcggcaaggc	catgtacgcc	1080
cccccatcc	gcggccagat	ccgctgcagc	agcaacatca	ccggcctgct	gctgacccgc	1140
gacggcggca	aggagatcag	caacaccacc	gagatettee	gccccggggg	cggcgacatg	1200
cgcgacaact	ggcgcagcga	gctgtacaag	tacaaggtgg	tgaagatcga	gcccctgggc	1260
	ccaaggccaa					
	tcctgggctt					
accctgaccg	tgcaggcccg	ccagctgctg	agcggcatcg	tgcagcagca	gaacaacctg	1440
ctgcgcgcca	tcgaggccca	gcagcacctg	ctgcagctga	ccgtgtgggg	catcaagcag	1500
	gcgtgctggc					1560
tggggctgca	gcggcaagct	gatctgcacc	accgccgtgc	cctggaacgc	cagctggagc	1620
	tggaccagat					1680
gacaactaca	ccaacctgat	ctacaccctg	atcgaggaga	gccagaacca	gcaggagaag	1740
	agctgctgga					
agcaagtggc	tgtggtacat	caagatcttc	atcatgatcg	tgggcggcct	ggtgggcctg	1860
	tcaccgtgct					1920
	cccgcttccc					1980
gagggcggcg	agcgcgaccg					2040
atctgggacg	acctgcgcag					2100
ctgatcgccg	cccgcatcgt	ggagctgctg	ggccgccgcg	gctgggaggc	cctgaagtac	2160
	tgctgcagta					
	ccatcgccgt					2280
atcggccgcg	ccttcctgca	catcccccgc	cgcatccgcc	agggcttcga	gcgcgccctg	2340
ctgtaactcg	ag					2352